DOCUMENT RESUME

ED 080 549 TM 003 054

TITLE	Assessing the Current Educational Congeries; The Annual Western Regional Conference on Testing Problems (16th, San Francisco, California, May 5, 1967).
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	5 May 67
NOTE	69p.
AVAILABLE FROM	Educational Testing Service, Princeton, New Jersey 08540 ($1.50)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	College Credits; *Conference Reports; *Educational Change; Educational Innovation; Equivalency Tests; Performance Criteria; *Program Evaluation; *Student Evaluation; *Testing Problems

ABSTRACT

The 1967 meeting of the Western Regional Conference on Testing Problems dealt with evaluation and educational change. The following speeches were presented: (1) "The Atypical Student" by Carole A. Leland; (2) "Citizenship Assessment" by Vincent N. Campbell; (3) "Educational Criteria and the Cheshire-Cat" by Scarvia B. Anderson; (4) "Are Educational Researchers Really Prepared to Evaluate Educational Programs?" by Rodney W. Skager; and (5) "Responsibility in Innovation" by Leland B. Newcomer. (KM)

ED 008 024

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

WESTERN REGIONAL CONFERENCE ON TESTING PROBLEMS

PROCEEDINGS

1967

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Dorothy Urban
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

EDUCATIONAL TESTING SERVICE

ED 000347

The Sixteenth Annual

Western Regional Conference on

Testing Problems

Assessing The Current Educational Congeries

May 5, 1967 · Hilton Inn

San Francisco International Airport

John S. Helmick, *Chairman*

EDUCATIONAL TESTING SERVICE
Princeton, New Jersey · Berkeley, California

Copyright © 1967 by Educational Testing Service
All Rights Reserved

$1.50 per copy

Preface

Although the theme of the 1967 Western Regional Conference on Testing Problems refers to a "congeries," the papers as presented belied that term. They turned out to be much more parts of a unified whole than a mere aggregation.

The program began with Carole Leland discussing the ways in which measurement of the individual can be used to enhance the educational process with special emphasis on credit by examination at the college level. Vincent Campbell then described assessment, not of individuals, but of the total educational effort as a necessary step in knowing what we are now doing so that we can make meaningful changes if required. In all of this, the problem of the criterion is critical and Scarvia Anderson called attention to this and described approaches to better statements of criteria. Rodney Skager then posed some interesting questions concerning the limitations and strengths of educational researchers in problems of evaluation. The final paper by Leland Newcomer provided a fitting close to the conference by drawing attention to the role of the administrator as a responsible person in innovation and by pointing out the dangers inherent in letting what has been innovation become routine.

Taken together, the various papers made a challenging but still hopeful picture of the role that measurement and evaluation can play in insuring that educational change becomes educational improvement.

JOHN S. HELMICK, *Chairman*

Contents

Morning Program

THE ATYPICAL STUDENT 1

Carole A. Leland, Assistant Director, Western Regional Office
College Entrance Examination Board
Palo Alto, California

CITIZENSHIP ASSESSMENT 13

Vincent N. Campbell, Associate Director
Social and Educational Research Program
American Institute for Research
Palo Alto, California

EDUCATIONAL CRITERIA AND THE CHESHIRE-CAT 23

Scarvia B. Anderson, Director, Curriculum Studies
Educational Testing Service
Princeton, New Jersey

Afternoon Program

ARE EDUCATIONAL RESEARCHERS REALLY PREPARED TO EVALUATE EDUCATIONAL PROGRAMS? 32

Rodney W. Skager, Assistant Professor
University of California
Los Angeles, California

RESPONSIBILITY IN INNOVATION 48

Leland B. Newcomer, Superintendent
Newport-Mesa Unified School District
Newport Beach, California

The Sixteenth Annual Western Regional Conference on Testing Problems

The sixteenth annual meeting of the Western Regional Conference on Testing Problems was convened at 9:15 a.m., Friday, May 5, 1967 in the Hilton Inn at the San Francisco International Airport. John S. Helmick, vice-president of Educational Testing Service, Princeton, New Jersey, presided as chairman.

The Atypical Student: His Alternatives and Ours

CAROLE A. LELAND

Like many of you, I have spent a good deal of time in the last few weeks in the educator's version of spring training—the myriad of conferences, annual meetings and hastily assembled committee sessions which constitute our warm-up for a new season. We watch our professional colleagues flex their muscles, so to speak, after the lonely, cold winter. We observe the scramble for new players and the fanfare of illustrious signings. And we note, often with dismay, that the leagues are growing bigger, the fans noisier, and the benches harder. Given this exposure it is extremely difficult to visualize how I might avoid saying the same things I have heard, thus continuing that endless cycle of issues and questions which seem to characterize our educational enterprise. (I might say too, just to stretch this analogy a little bit, that I am not quite sure anymore when spring training ends and the season begins.)

In addition to this possibility for repetition let me admit from the outset, something which could become increasingly obvious, that in all my collegiate experiences I have never had a course in public

speaking. I have consulted with a friend who has, however, and I do know that one might classify speeches in four categories: argumentative, persuasive, informative and entertaining. Now, it seemed logical to me that if I covered all four in one speech I might compensate for my lack of formal training.

The entertaining function I prefer to dispense with first and quickly. At this early hour I am sure you wouldn't expect more. Two stories seem appropriate to me, one which I find relevant to students and one which seems to have some relevance to the learning process. The first concerns two cows grazing in the meadow when a dairy truck passes nearby. The side of the truck announces grandly, "Grade A Milk, fortified with Vitamin D, B, and B_2, enriched, homogenized, pasteurized, lower in fat, higher in protein than regular milk." One cow turned to the other and said, "Sort of makes you feel inadequate, doesn't it?"

The second story concerns a woman who took and successfully completed the Red Cross's course in first aid. Some months later she was moved to write the national director of the Red Cross in Washington. Her letter went like this:

> Dear Sir:
>
> I am extremely grateful for having had this opportunity to take the Red Cross course in first aid and I want to tell you how meaningful it has been for me. Very recently we had a horrible accident near our home, with bodes strewn all over, wreckage all about and because I had had your course in first aid I knew exactly what to do. I put my head between my legs and I didn't faint.

The presumptuous title which appears on your program is exactly what you might expect, the result of pre-planning when a speaker agrees to anything that is several weeks off, and the person doing the program accepts anything in order to get it to the press. For who among us today can distinguish confidently between the atypical and the typical student? Most of us are hard put to distinguish between male and female, student and nonstudent, and faculty and administrators on the college campus.

Nonetheless, I prefer not to abandon the topic and only wish to expand it a bit to read, "The Atypical Student—*His* Alternatives and *Ours*."

What I have in mind is to argue a common point of view about American Higher Education, to persuade you to delineate more

carefully the uniqueness of the students you deal with and in the process to recognize individuals who may be excluded unjustly from our colleges and universities, and finally, to inform you of the College Board's concern for providing alternatives to a diversity of individuals who seek collegiate instruction.

Now the argument is neither original nor complicated. For all its magnitude and magnificence, American higher education has responded lethargically to the challenges of an increasingly diverse college population. This is hardly "kicking a man when he's down," for higher education has never been so "up." With larger numbers of students, greater interest and support from large segments of society and unbelievable concentrations of brain power in their students and faculties, the 2,000 odd colleges and universities of this country are, to put it mildly, "loaded." To be sure, they are loaded with problems as well as potentials, but to rationalize the dormancy of the latter in the face of the former is sheer delusion. Admittedly, too, some institutions are making efforts to accommodate all kinds of students whose motives, dispositions and capacities differ, sometimes markedly. We have welcomed the residential complexes of Santa Cruz or Michigan State, the technological daring of Oakland Community College and the overseas campuses of numerous schools as evidences of a willingness to meet the challenges of diversity. But, in the main, when viewed collectively, the colleges and universities have exhibited little passion and ingenuity in engaging large numbers of students in the learning process. As a matter of fact, large numbers of students have been repelled and some have never been attracted to the campuses at all.

Martin Meyerson, writing in the recent volume *The Contemporary University: USA*, suggests that in our culture we tend to cater to majority tastes and choices in most fields, and higher education is no exception. He opines:

> Yet, if the number of students in higher education increases as rapidly as we expect, in about 15 years the present number of students may be only half the total. With such vast numbers of college and university students, minority interests of student subcultures could fill many campuses. One of the prerequisites for high quality will be the provision of more educational diversity for these minority interests.

And he concludes: "A new academic *ethos* of diversity and yet

community will require far more spontaneity in organization than educational institutions have commonly exhibited."[1]

Mr. Meyerson conveys the sense of diversity but too little of the urgency some of us feel at the present time. Earlier I side-stepped the matter of "atypical," implying that our sense of the typical has been buried in the flutter and flurry of rapidly changing and growing college populations. Certainly that may be true but such caution hardly gets us into a position of intellectual valor and challenge. So let me suggest that one way to view the atypical student is to consider him as someone who doesn't quite "fit" the traditional pattern of undergraduate education in most of our colleges and universities. Of course, this forces us to view a good many students as atypical. And that's the point.

If I am to persuade you to recognize and to cultivate the atypical let me select a few examples among the many available to us. (And let me just digress a moment here to say that probably the most disappointing part of any speech is the necessity of selecting examples and frequently leaving behind the one that might have been the most exciting and the most dramatic.) The possibilities which could challenge our institutional sensibilities are numerous—the drop-outs, the uncommitted, the alienated, the "retread" to name but a few. Let me select three groups, however, which are perhaps familiar and more common to all our experiences but which should provoke, if they haven't already, both our concern and our ingenuity. Each group, either by its own urgings, or in the light of its unique situation, speaks to the inadequacies of the institutions of higher learning.

The student activists deserve mention, not because they haven't received our attention (they are far too vociferous and we are far too curious to have missed them), but because they constitute one of the few groups of students we know much about. Kenneth Keniston calls attention to our vast ignorance of students in the following quotation which is taken again from *The Contemporary University: USA:*

> To be sure, everyone acknowledges that without students there can be no university; and so too, "education" is widely admitted to be one of the functions of a university. Yet the characteristics of students —the fact that they have commitments, aspirations, values, dreams, needs, psyches and perhaps souls even *before* being admitted to college—are largely ignored in the concentration on more easily describable features of the university. To many administrators and to

> some faculty members, students remain a kind of unleavened lump to be molded by the university, blank cards on which education will punch imperishable information, shapeless ingots to be pressed into useful forms by "the college experience."[2]

In contrast to this paucity of knowledge about students generally, we know a good deal about the student activists, particularly those in the 1964 FSM and those students who demonstrated at Chicago and Pennsylvania State. In a monograph prepared for the U.S. Office of Education, Joseph Katz emphasizes that in numerous studies the major findings are essentially convergent. And I am sure you will agree with all of us in research that this is amazing. Katz summarizes these data in the following way: The activist tends to come from homes with parents whose incomes, occupational status and education are higher than parents of non-activists and the parents are likely to be politically more liberal, more permissive in child-rearing and to have had closer affective relationships with their children. In terms of academic aptitude, performances and attitudes, the students themselves scored significantly higher in verbal, not in mathematical aptitude and their grade point averages were significantly higher than those of non-activists. They scored higher on scales measuring theoretical orientation, liking for reflective thought, diversity of interests and estheticism. The activists reported themselves more often than non-activists as having been influenced by ideas presented in courses and by teachers. Measures of personality characteristics and values, among other things, revealed the activists to be more flexible, tolerant and realistic, less dependent upon authority, rules or rituals for managing social relationships, less judgmental and tending to express impulses more freely either in conscious thought or in overt action. In his description of the activists, Katz suggests:

> That the activists are recruited particularly from the intellectually able and interested students may well be explained by the observation that these are students who care for the intellectual values implied in the university, who have been stimulated by the pursuit of truth and expression of heterodox ideas in their courses, who now want an extension and deepening of these experiences and who would like to relate ideas and theories to their own lives and to the improvement of society around them.[3]

We continue to observe and to assess the impact of the student activists, but above all we recognize in their demands for educational

reforms a desire for our concern and commitment to them as individuals. However atypical their methodologies, their demonstrations should call attention to the rigid pathways we have constructed throughout the undergraduate curriculum.

For a second group of students, the collegiate experience is often equally frustrating and straining, but there have been too few of them to constitute more than a silent minority on the campus and as individuals we know little about them. For the talented student in the arts, higher education has made only token gestures of awareness. To be sure, we see more art exhibits, more musical productions and dramatic events and more artists in residence on some campuses. And some colleges, Bennington and Sarah Lawrence for example, place a premium on the place of arts and artists within the society and within the gates of the college.

But, in a position paper prepared for the National Foundation of the Arts and Humanities, Lewis Mayhew dramatically presents the argument that while the fine and performing arts seem ascendant in American society, American higher education has yet to demonstrate that the arts should feature strongly in the life of each college and university. Such a demonstration, he suggests, involves both providing a comfortable setting for the artist within a college and ensuring that students with interests or talents in the arts can enter the university without undue guilt or without jeopardizing their primary interests. Professor Mayhew bases his somewhat pessimistic assessment on three indications:

1. The fact that the normal predictors of academic success, like high school rank and performance on aptitude tests, used to screen students for higher education bear little or no relationship either to performance in the arts or to creativity in the arts, and interest or talent in the arts is rarely among published admission criteria.
2. That the collegiate curricula themselves reflect little concern for the performing arts and few college honors programs are designed in the arts, and
3. The continuing posture of most collegiate faculty that the arts are frills in comparison with other subjects.[4]

In like fashion, Dr. Benjamin Steigman, for 22 years principal of New York's High School of Music and Art, makes the point that anything like parity for high school music and arts courses is out of the question. He states, "The college catalogs appear to be liberal. In most of them there is this reassurance, 'A candidate whose record

shows special ability, but who lacks the required academic credits, may present his case for consideration by the Committee on Admissions.' With four or five times as many applicants to the prestige colleges as they can admit, what student would add to his hazards by lacking the required academic credits? Just how would his special ability be evaluated? No, he had better stick to those courses he knows are acceptable."[5] And, I might add that Dr. Steigman substantiates his position with the report that out of 110 well known colleges, 90 percent admitted they would allow no substitution of music and art for either math or foreign languages.

The student of the arts, like the student activist, is atypical in most of our colleges and universities because too often his "fit" to the traditional curriculum and pattern of campus life forces compromises upon him and frequently forces him off the campus. In a 1961 address to the Association of Graduate Schools, W. McNeil Lowry called for a "radical shift in atmosphere" surrounding students considered potential artists. He elaborated that the essence of this atmosphere is "the artist's acceptance of concentration, even of distortion, as a way of life, a way of life that in many ways is completely antithetical to the ideal objective of a liberal and humane education."[6] Despite some attention in individual institutions and by some recognized leaders, when change is measured in such terms as Mr. Lowry's, our collegiate enterprise has shown relatively little concern for individuals in the arts or for those who might be in the arts if the barriers were not so great.

Time prevents more than a passing comment or two about a third group of students, some of whom we recognize on our campuses, others we never see. Increasingly since the late 19th century the adult student has struggled for attention in colleges and universities primarily developed for educating the young. That struggle has not been easy, nor has it been won. The 1961–62 National Opinion Research Center survey revealed that some 25 million adults were engaged in some educational course or activity. However, churches and synagogues served 3 million of these students, compared to 2.6 million in college and university classes and 8 million were involved in independent study.[7] A majority of the adult students have enrolled in traditionally vocational courses.

On some campuses the adult student has been welcomed, encouraged and his purposes in approaching the institution have been respected. Special degree programs at schools like Brooklyn, Queens,

Syracuse and Oklahoma have accredited previous learning and experience, provided flexibility in course schedules and ensured the involvement of students with top-level faculty. Experiences with adults have proved not only satisfying but invigorating for many schools. In a progress report of the Center for Continuing Education and Community Studies at Sarah Lawrence College, B. J. Loewenberg reports: "The experience of the Center has justified its initial purposes. The performance of mature women returning to undergraduate work and professional study has confirmed our faith in their undiminished intellectual vigor." [8]

Swift appraisal of the adult's situation with respect to higher education suggests progress and encouragement. But more careful scrutiny underscores that whatever responses we have made to the adults who seek our attention have been minimal and sometimes superficial. We know very little about adults as students; we have been largely intolerant of their necessity for institutional flexibility and we seldom view the adult as an individual whose previous experiences have relevance or respectibility in the academic world. If you doubt these accusations, talk to a young housewife who has completed three years of college and wants to complete her education at a nearby university which refuses part time students, or listen to the military man who seeks a higher degree and is turned away because he has reached his 45th birthday.

The question really isn't who is typical and who is not. The point is that the educational establishment is full of individuals who for one reason or another don't exactly fit our carefully devised four-year curricula. We may also do great injustice to that seemingly large group of students who adapt to our mold dutifully or resignedly.

In this context of diversity, realized in some quarters, potential in others, the College Board has been trying to make some headway. A year ago you heard Richard Pearson, the President of the College Board make this proposition:

> Any other individual (and here he was speaking of those individuals who do not progress directly from high school to a two or four year college), of whatever age, should expect opportunities to demonstrate through independent study or other off-campus learning experience that he is qualified for formal study at the college level.[9]

Let me describe the extent of the Board's commitment to this principle. Seemingly this informative aspect of my comments should be

the easiest. For here I can rest with confidence on the wisdom and energies of those College Board and ETS staff members who have made the College Level Examination Program a reality. But reflection about the scope and daring of the venture alone could consume all the time allotted for this presentation. Thus the information I offer is sketchy and hardly does justice to the intentions of the College Entrance Examination Board which sponsors the program.

CLEP, the inevitable acronym for the College Level Examination Program is essentially a national system of placement and credit by examination. I stress this *essentially* for two reasons: 1, the explanation abbreviates considerably the diverse uses of the program and 2, it may generate unwarranted suspicions about its purposes. The major objectives of the Program attest to its real and potential diversity:

1. To provide a national program of examinations that can be used to evaluate nontraditional college-level learnings, specifically including independent study and correspondence work;
2. To stimulate colleges and universities to become more aware of the need for and the possibilities and problems of credit by examination;
3. To enable colleges and universities to develop appropriate procedures for the placement, accreditation and admission of transfer students;
4. To provide colleges and universities with a means by which to evaluate their programs and their students' achievement;
5. To assist adults who wish to continue their education in order to meet licensing requirements or qualify for higher positions.

In many ways, CLEP's beginning, which dates back to 1965 when the College Board's Council on College Level Examinations was appointed and began the Program's development, has been auspicious. First, the Program is already operational because it rests on the College Comprehensive Tests, initiated and developed by ETS, now sponsored by the College Board as the first phase of CLEP. The original concern and investment by ETS in no small measure accounts for the positive response to the Program. Second, CLEP has enjoyed the generous support and financial assistance of the Carnegie Corporation. In addition, it has the endorsement of the Federation of Regional Accrediting Commissions and the Commission on Accreditation of Service Experiences. Finally, the Program begins with strong leadership by some colleges and universities and

with at least attentiveness on the part of a great many more institutions.

Today CLEP consists of two types of examinations. The General Examinations which are designed to provide a comprehensive measure of undergraduate achievement in five basic areas of the liberal arts—English composition, humanities, mathematics, natural sciences and social sciences (history). They are intended to measure an individual's knowledge of fundamental facts and concepts, ability to perceive relationships and understanding of basic principles. The five examinations taken as a battery include those subjects typically required of general education students in the first two years of college. The Subject Examinations, and there are currently 11 of them, are end-of-course tests appropriate to courses taught in the undergraduates years on either a semester or yearly basis.

Faculty members from colleges and universities throughout the country serve the College Board well as the examining committees. With the assistance of test specialists from ETS they develop the examinations and perform the consistent review which ensures the appropriateness of the program to the collegiate enterprise.

The future of the College Level Examination Program rests heavily on three things: One, the vision and capacity of the College Board to expand the program when and where appropriate, while maintaining the quality of its initial offerings. And that expansion has already begun. By 1968 five new subject examinations will be available—marketing, money and banking, statistics, history of American education and educational psychology. Further, examinations will soon be available not only at institutions with which an individual contracts, but also at national testing centers, as is the case with the admissions testing program. Two, the future of CLEP depends upon the evaluation of the present use of the Program. Let me cite just four examples of current usage to indicate the diverse purposes which CLEP may serve and to illustrate the point that assessing intellectual achievement on a national basis does not necessarily stifle, threaten or usurp the prerogatives of individual institutions.

The University of Iowa, one of the country's major state universities, has recently authorized the use of the College Level Examination Program's General Examinations as an alternative way for students to meet graduation requirements. That is, students may meet requirements either by taking approved courses or by exami-

nation. Montgomery Junior College, a public institution in Maryland, uses the General Examinations to provide an indication of how well its graduating sophomores perform in comparison with the national sophomore normative data and to provide a counseling tool for working with their graduates. In the School of General Studies at Queens College of the University of New York, the General Examinations are being used to assist with the placement, screening and admission of adult students. San Diego State College participates in CLEP through its use of the examinations to insure basic levels of achievement in the liberal arts area of secondary education candidates who seek to enroll in student teaching. In these examples, and others, the institutions decide not only how the exams are to be used, with or without credit provisions, but what standards of performance are acceptable.

Still, the future of CLEP depends to a great extent on the willingness of colleges and universities to react to the Program and to shape its development. The Program at times seems embarrassingly simple. It presupposes no particular *pattern* of preparation nor expects any common background. It provides an alternative for recognizing and rewarding achievement and for encouraging students to continue an education which at times doesn't appear to conform to traditional patterns.

I began with Spring training; let me close with three quick pitches:

1. If we continue to operate in a vast ignorance of the true diversity of students and potential students in this country we engage in shallow investments and superficial concerns which degrade the entire educational process.

2. The alternatives available to the student, typical or atypical, reflect little comprehension of his individuality and precious little regard for what learning is all about.

3. Programs like the College Level Examination Program are designed to serve human beings and diversity. If they solidify into one more obstacle to an individual's self-realization, we aren't doing our job and neither are you.

REFERENCES

1. Meyerson Martin, "The Ethos of the American College Student: Beyond the Protests," in Morison, Robert S., *The Contemporary University: USA*, Boston, The Houghton Mifflin Company, 1966, 289, 291.
2. Kenniston, Kenneth, "The Faces in the Lecture Room," Op. cit., p. 322.
3. Katz, Joseph, "The Student Activists: Rights, Needs and Powers of Under-

graduates," Monograph prepared for the United States Office of Education, 1967.
4. Mayhew, Lewis B., "The Arts and Access to American Higher Education," Position Paper prepared for the National Foundation of the Arts and Humanities, 1966.
5. Steigman, Benjamin, *Accent on Talent,* Detroit, Wayne State University Press, 1964, 167–68.
6. Lowry, W. McNeil, "The Arts" in Morison, Robert S., *The Contemporary University,* Boston, The Houghton Mifflin Company, 1966, 213.
7. Johnstone, John W. C. and Rivera, Ramon J., *Volunteers for Learning,* Chicago, Aldine Publishing Company, 1965, Chapter 3.
8. Loewenberg, B. J., *Work in Progress, Report III,* Sarah Lawrence College Center for Continuing Education and Community Studies, 1966, p. 29.
9. Pearson, Richard, "Access to Higher Education: Implications for Future Planning," in Educational Testing Service, Western Regional Conference on Testing Problems, Proceedings 1966, p. 12, Berkeley, Calif.

Citizenship Assessment

VINCENT N. CAMPBELL

I think I win the prize for the least inspiring title on the program. But let me try to compensate slightly for that by saying that I want to discuss, not just citizenship assessment, but the national assessment effort as a whole. I think more of you are familiar with the whole enterprise than you are with any particular subject matter or particular part of the assessment effort.

About $40 billion are invested annually in American education, and a good question is: What kind of evidence do we have that it is achieving what we really want it to achieve? This is an especially appropriate question in view of the fact that we are now in the midst of a flurry of educational innovation, and there seems to be a general feeling that things need changing. Teachers, school administrators and other people concerned may have pretty good intuitive feelings about what needs improvement, but in view of the size of the enterprise and the importance of the outcome, it seems risky to rely on these intuitive feelings as a way of deciding whether education is achieving what it is supposed to be or not.

As I see it, the purpose of the national assessment is to measure specifically what important achievements are actually being reached and which ones we think that are important are not being reached.

The national assessment is an idea which has not yet gained full acceptance, as many of you are aware. The whole program is divided into two phases. In the first or exploratory phase, now nearly completed, the objectives of assessment were developed and the first draft instruments to measure these objectives were prepared. After these are revised and reviewed thoroughly, if it appears that the assessment is ready to proceed, students from all over the

country will be sampled and measures of their achievement in eight major areas will be obtained. These areas are: citizenship, vocational education, social studies, the sciences, mathematics, reading, music and art.

Many educators have asked why such an elaborate enterprise is needed to measure the achievements of education. Why not look at ordinary classroom testing results or use existing standardized tests. The answers are straightforward. Ordinary teacher made tests vary a great deal from school to school, both in content and in quality and they are seldom based upon a careful specification of the objectives which students are meant to achieve. Another objection is that many tests are predominantly tests of memory. When we ask teachers what student achievements are most worthwhile they claim that memorization is not that important, yet memory is mainly what they are testing for. Right there is a mismatch that needs some attention, and maybe the national assessment will help point this out.

As for standardized tests, the main failing is that these are not designed to assess the achievement of a group on specific important objectives. They are designed rather to discriminate among individuals. Important achievements which nearly everyone masters, or which hardly anyone does, are purposely left out. But one purpose of the national assessment is to decide what objectives are being achieved by nearly everyone and what objectives are being achieved by very few people, as well as those which are being achieved by about half of the students in the country.

Existing tests and data, then, were considered inadequate for a national assessment. The first step in the plan therefore was to design objectives and develop new measuring instruments to measure achievement of these objectives. The contractors which have been developing these instruments are the Educational Testing Service, the American Institutes for Research, the Pyschological Corporation and Science Research Associates.

Now let me describe briefly what I see to be the main features of the national assessment as a whole. The main purpose is to obtain accurate information on the achievements of the nation as a whole and major subgroups within the nation as a basis for better planning of public policy and research.

Second, only those achievements agreed to be important by scholars, schools and laymen are being assessed. There is no implication here that achievements not being assessed are not important.

It is solely that in order to make the best use of the money available the idea is to assess objectives which all three groups; schools, scholars and laymen agree are important.

Third, each student sampled will take only a small fraction of the measures. No score for an individual is wanted, and the total number of measures required to thoroughly assess all these specific objectives is very great. Perhaps 50 or 100 hours of assessment in all, or more. It would be unfeasible to expect one student to take all these measures. Therefore each student's participation will be limited to one hour or less.

The target age levels for the assessment are ages 9, 13, 17 and adult. A new assessment is planned to take place every few years and the objectives and measures will be updated and revised before each new assessment. The purpose of repeated assessments, of course, is to give an idea of the progress that is being made over the long course.

And last, the results will be reported in a form that is understandable to laymen rather than as abstract test scores. For example, we are not trying to get a citizenship score of 40.6 with a standard deviation of something or other for a person or even for the nation as a whole. Rather, for each measure of an objective we want to know what percent of the population is achieving it. It might be reported, for example, that 10 percent of the 17-year-old urban boys in the Midwest have written a letter to a public figure about some specific issue. In such terms we feel we can communicate to all the people who are concerned about education. They don't have to guess at what went into a test score. They can see the measures, or at least samples of the measures on which results are reported.

Several people have asked why citizenship is included in the assessment program since it is not a traditional discipline in the educational enterprise. The reason is that many important goals that education could achieve are not included in the traditional subject matter disciplines. People agree that citizenship is a worthwhile thing, but it tends to get relegated to limbo, the only vestige being a global rating by teachers at the end of six weeks. It really receives very little attention in the educational enterprise. With the exception of a few schools, there is not a strong educational effort focused on citizenship.

Also there has been very little specific assessment on a broad scale of civic achievement either in education or outside. There has been

even less in the basic disciplines of mathematics, science and social studies. The few careful efforts at citizenship measurement that have been made have occurred in experimental civic education programs such as those in Detroit and Kansas.

To many people evaluating citizenship probably calls to mind giving a student some sort of over-all rating. I want to make it clear from the start that that is not the approach here at all. We are not saying how good a citizen Johnny is on the whole. We are instead constructing measures of how well specific behavioral objectives in the citizenship area are being achieved by large population groups.

The first stage of our procedure was to decide what the most important citizenship objectives are—what do schools, scholars and laymen want students of various ages to accomplish? Next, we developed specific exercises to measure the extent to which each objective is being achieved. For each objective a rationale was written describing why certain kinds of measures seemed appropriate and why other kinds seemed inadequate. The purpose of the rationale was to set forth our reasoning for inspection, not only by ourselves, but by other people who were concerned with the quality of the measures. Also if a measure wasn't working out another could be substituted for it on the basis of the rationale.

Our standard for choosing citizenship objectives was their importance to consensual civic goals in our society. Whether the objective is being achieved well or poorly, now, and how much schools, home, church or other institutions have contributed to the present level of achievement, were not considered at all. We did not examine the current curricula of schools in the area of citizenship. Nor did we choose the objectives which would be easiest to measure, though it is a very tempting thing to "search where the light is best" when you are working in a difficult area. Consequently, importance to society, as agreed upon by teachers, scholars and laymen, was the sole basis for selecting objectives.

In order to keep to the heart of the matter, some areas related to citizenship were not included in our objectives. One class of objectives excluded was those which are clearly more central to some other subject matter area being assessed. For example, "knowledge of ways of life in other cultures," we assumed, would be covered in social studies. "Religion," "living a rich, full life" and "vocational competence" which are important in a well-rounded life were judged to be only tangentially related to citizenship. So by logical

sequence, only those aspects which were judged to be of particular civic relevance were included. For example, we measure freedom of religion as a constitutional right because it was of particular civic relevance, but religion as a whole was not covered.

We started the process of choosing objectives by reviewing previous lists of citizenship objectives and boiling these down to one parsimonious but comprehensive list of about twenty objectives. The failing of many writers of educational objectives in the past, we thought, had been that the objectives were confined to glowing generalizations which did not have much effect on anything. In order to get at the crucial specifics, we took two steps:

First, outstanding local teachers familiar with each target age group worked for weeks with our staff to break down each general objective into the most germane behaviors appropriate as goals for a given age group. They drew on examples from their own experience and compiled a list of very specific behaviors beneath each general objective.

And second, students and adults in each age group were asked to recall and describe outstanding citizens of their acquaintance and to recollect specific incidents and descriptions. These incidents and descriptions, about a thousand in all, were used to check the completeness of our initial list of objectives.

The revised list of objectives, each broken down into important behaviors, was then worked over for three days by a panel of national leaders in citizenship, education and related social sciences. Persons in varying community roles also reviewed our objectives and made comments, and the objectives were further revised. Some of these community consultants were public and private school administrators, counselors and teachers, a judge, a county planner, social scientists, labor and business leaders and others. The final list was submitted to the exploratory committee directing the whole assessment program, which is chaired by Ralph Tyler and has Jack Merwin as its current staff director. The committee held meetings with panels that were made up of concerned laymen from all regions of the country for the purpose of critically reviewing the objectives submitted by the contractors in all subject matter areas. A few wording changes were recommended by the laymen, but they did not recommend elimination of any of the citizenship objectives which are listed in abbreviated form below for seventeen year olds.

CITIZENSHIP OBJECTIVES

Acts of good citizenship by 17-year-olds:

I. *Show concern for the welfare and dignity of others.*

They respect the worth and individuality of all persons regardless of religion, beliefs, race, or other social or personal characteristics. They weigh the effects of their own actions on the health and welfare of others. They seek to better the living conditions of less fortunate people, including people who are denied equal socio-economic opportunity. They help other individuals in need. They are loyal to country, to mankind, and to other groups whose values they share.

II. *Support constitutional rights and liberties of all individuals.*

They recognize when constitutional rights are being properly exercised or denied and defend those rights, regardless of the victim's unpopularity. They understand the values of our traditional freedoms and due process of law.

III. *Help maintain law and order.*

They understand the need for law and try to keep informed of the law. They comply with the law and school regulations, or when they think a rule unjust, they oppose it by lawful democratic means. Their own behavior is ethical, and they encourage ethical behavior in others.

IV. *Know the main structure and functions of our governments.*

They understand the reasons for having government and the constitutional source and limitation of our governments. They know the main powers of each branch of government, and the relations among levels in our federal system. They value political opposition and see the need for having a variety of interest groups. They recognize the importance of citizen participation in government and the many ways in which a citizen can influence public policy.

V. *Seek school and community improvement through active democratic participation.*

They believe that each person's civic behavior is important and each should do his part to help solve society's problems. They take an active part in groups which seek to better the school or community. They help achieve group goals by cooperating with other group members, and they support democratic procedures in group meetings.

VI. *Understand problems of international relations.*

They seek world peace but are aware of the many causes of international conflict and of dangers to national security. They favor constructive efforts to increase understanding and accommodation among conflicting nations.

VII. *Support rationality in communication, thought, and action on school and community problems.*

They seek out and critically interpret information of civic importance. They try to understand and weigh the merits of alternative

viewpoints and actions. They are effective problem-savers and planners. They support education and free communication, and communicate their own views clearly and honestly.

VIII. *Take responsibility for their own personal development and excellence.*

They make maximum use of opportunities to further their own education and prepare for adult roles. They show initiative and diligence, and fulfill voluntary commitments.

IX. *Help and respect their own families.*

They respect their parents' wishes and the views of all family members. They aid the civic development of younger siblings and help with home duties. They discuss and relate peer group influences to family values.

The nine main objectives represent "clusters" of related important civic behaviors. The order of listing the nine is not hierarchical and has no special meaning. The list is quite similar for the other age levels. However, as each objective is broken down into more specific behaviors they are different at the different age levels. It is only at the general level that there is considerable similarity. We expect a more complete description of the objectives to be available in published form in the not-too-distant future after the exploratory committee has had a chance to pass on them.

Please keep in mind that these objectives are intended solely as a working guide for the difficult task of assessment. We are not prescribing how an individual citizen should behave or how he should spend his day. In fact, no one person could be expected to perform all the specific behaviors we have included. Some achievements, such as seeking public office or raising children well will naturally only be attempted by certain citizens. But since the assessment is intended to describe the achievements of a wide population of citizens, not individual persons, this presents no problem.

Let us now turn to the achievement measurement. We started, as I explained, by writing rationales in order to make our reasoning in developing these measures explicit and reviewable. After we had prepared the draft measures they were submitted to two kinds of tests. First they were tried out with a few students, the number of students depending on the nature of the measure. And second, they were reviewed again by committees of laymen from around the country to see to what extent they were objectionable.

Most of the measures which we have developed in the area of citizenship are intended to determine what citizens do or have done

rather than whether they know what civic behavior is proper or not. For some objectives, we do seek knowledge as the achievement sought, but often where knowledge is assessed it is taken as an indicator of past behavior rather than as knowledge for its own sake. For example, if we ask students to name a favorite book or game of their younger brothers or sisters, we are not valuing this knowledge in itself, but using it solely as an indication of how much attention they give to younger family members.

We think that most important citizenship objectives cannot be assessed adequately by paper-and-pencil techniques alone. Interviews will be used extensively, especially with nine-year-olds. In fact, in the citizenship area we have not proposed *any* paper-and-pencil measures for nine-year-olds. In order to assess reactions to spoken messages, we are having some of the older students listen to tape recordings and then respond to these communications in certain ways. Another type of measure we are using is recall by teachers and students of specific incidents of certain kinds such as helping other students and school assignments completed or not completed.

Direct observation of behavior in groups is an expensive measurement technique, but it was our conclusion that it is the only valid way to assess certain objectives in the area of citizenship. For example, we could think of no less expensive way of measuring "effective cooperation with others to achieve a group goal" than by observing people in their interaction and recording certain specific behaviors observed. Another objective of this sort was the willingness to express one's views in public. A third objective was the respectful treatment of other individuals. For each school age level we therefore contrived one or two standardized group situations in which these and other behaviors could be observed. Any measures which we have developed involving direct observation of behavior, judging by our tryouts to date, will require the use of a checklist to simplify the coding of the observations and at least an hour or two of training for each observer.

The tryouts are now underway for all subject matter areas in the national assessment, and the purpose of these tryouts is mainly to establish administrative feasibility. To some extent, the tryouts also help us in detecting our mistaken premises and other weaknesses in the measures. But mainly we want to know how the measures should be administered and who should do it. That is, what things can teachers in the schools themselves handle adequately and which

tasks require more expert administration. Also, how should the measures be packaged, that is, how many different subject matter areas or formats can a student respond to adequately in a half an hour.

Late in 1967 more extensive tryouts will be conducted to determine whether we are truly sampling all levels of difficulty. That is, are we getting measures of what 90 percent of the students can do and what only 10 percent can do, as well as of those achievements reached by about half the age group.

These tryouts are not validity studies in the ordinary sense. I am not sure whether there is any group on which you could validate gains for measures of this sort. The main criterion for the acceptance of any measure is, whether on inspection of the measure itself and the rationale for its use, one is convinced that the measure validly measures achievement of the objective it is intended to assess.

Before closing I should mention something of the controversy that has surrounded the national assessment. "Fear of national control" seems to be the theme brought up most. I have tried to understand this somewhat, but I must confess that I don't, and this is certainly not because I favor national control. On the contrary I am for local control of education. I am not even for state-adopted textbooks, as we have in California. The purpose of the national assessment is to gather information about the achievement of objectives. How this information is used is entirely up to the person who is using it. It seems to me that there is less reason for teachers to prepare a student to do well on a national assessment, since the chances are one in a thousand that the student will be involved in the assessment anyway, than there is for a teacher to prepare students to do well on, say, college entrance examinations which have considerably more effect on the individual lives of students. It doesn't seem to me that educators in general are exerting great effort to adapt their curriculum to the tests that now exist. So I don't understand exactly why it should be expected that the national assessment, which will not even characterize any individual student, should be feared on these grounds.

Another point that I would like to make is that in future national assessments these objectives will be updated so one need not fear that the objectives are going to become archaic and that a fixed set of measures is somehow going to determine the future course of education. Obsolescence does not seem to be a valid criticism of the

plan. We hope that national assessment will, on the contrary, periodically provide a current, accurate appraisal of our society's aspirations and its degree of success in achieving them.

Educational Criteria and the Chesire-Cat

SCARVIA B. ANDERSON

You will recall that the baby had just turned into a pig—rather a handsome pig though it would have made a dreadfully ugly child—and trotted off into the wood. Alice was startled to see the Cheshire-Cat sitting in a tree.

After some preliminaries, she ventured to inquire, "Would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to," said the Cat.

"I don't much care where—" said Alice.

"Then it doesn't matter which way you go," said the Cat.

"—so long as I get *somewhere*," Alice added as an explanation.

"Oh, you're sure to do that," said the Cat, "if you only walk long enough."[1]

There are many Alices among educators and would-be evaluators of what goes on in education. We are seeking direction, but we are not sure (and sometimes say, defensively, we don't much care) where we want to go—and our explanations are less than satisfactory.

But perhaps we are not entirely to blame. The objectives that we seek and the criteria by which we will know whether we have attained them are as elusive as the Cheshire-Cat. They keep appearing and vanishing so suddenly that they make us giddy, as Alice was. Even when the Cat obligingly disappeared quite slowly, beginning with the end of the tail and ending with the grin, you will recall that the grin remained some time after the rest of it had gone. And that was quite disconcerting to Alice, just as it is disconcerting to us to

realize that we have sighted some element of our criterion but the rest still eludes us.

Of course, in talking about educational criteria I am also talking about educational objectives, because the two are inextricably intertwined to the degree that an adequate explication of the criteria makes almost unnecessary any separate statement of the objectives, and an inadequate explication of the objectives makes almost impossible any definition of the criteria.

In this mood of shifting some of the blame for the situation we are in from ourselves to those ornery rascals, the criteria, let us look at some of the reasons why the criteria make themselves so hard to get hold of.

First, as Henry Dyer from the ETS branch office in Princeton says, "criteria rely too much on the magic of words."[2] An objective pompously introduces itself as "learning how to live and work together."[3] Now by what criteria would we know whether children had learned to live and work together?

Dyer goes on to mention two other reasons why educational goals have been nonfunctional in the past, and I take "nonfunctional" to mean unproductive of adequate criteria. One is that there has been too little public participation in formulating them; the other is that there has been too great a readiness to suppose that they are already given and require only to be achieved.

A fourth source of problems arises out of the field with which many of us here are most closely associated: measurement. Some criteria appear momentarily and then disappear because there are no existing, immediately available means of measuring them. In some cases, this discourages us; in others, we grasp our old security blanket statement, "Many of the things I teach are intangible and cannot be evaluated."[4]

What can be worse is that, if there is an existing measurement instrument, we are inclined to assume that there is a worthwhile objective to go with it. Thus, because we just happen to have a test which purports to measure behavior X, we suddenly decide that student competence in X was something we had very much desired.

Those of us who talk of educational criteria are sometimes accused of having some of the characteristics of the criteria we criticize. We tend to speak in vague generalities. Lest I fall completely into that trap, let me describe some of our experiences in searching out one particular criterion which has proved especially difficult and elusive.[5]

We are not sure of the proper *name* for it, but it has been *called* "school readiness," and it *is* something which most preschool programs say they are trying to develop and most first grade teachers decry the lack of in some of their pupils. (We shall not concern ourselves right now with the facts that some psychologists question the viability of "increased readiness for public school" as a major goal of preschool stimulation programs[6] and that most teachers admit that their job is one of coping with all children who come to them, regardless of whether they are "ready.") "School readiness"—at any rate the *words* are with us. Tests have the words in their titles, teachers write the words on their reports about children, and sponsors of preschool programs list them in their proposals for funds.

We used two kinds of detection devices as we went hunting "school readiness." They are both rather old-fashioned. The first is the "critical incident" technique; the second, the group discussion. The first detection device was activated last fall with the help of a national sample of about 300 first grade teachers. They were asked to do the following things:

1. List all of the children in their classes and tell us when they were born and whether they were boys or girls.
2. Answer the following question about each child: As of this date, how ready do you think this child is for the first grade experience? —Ready in all respects? Ready in most respects? Ready in only a few respects? Not ready? (We had originally considered a three-point rather than a four-point rating: ready, not ready, and can't judge. But the group of teachers on whom we tried out the questionnaire indicated—very intelligently— that very few children were wholly ready or not ready. They were also unwilling to admit that they couldn't judge any child!)
3. For each child rated ready in all or most respects, give a specific example of one thing which he did or said which suggested that he *was* ready. Or, alternatively, for each child rated not ready or ready in only a few respects, give a specific example of one thing which the child did or said which suggested that he was *not* ready.

At a later date, teachers also supplied us with information about children's preschool experience and home background.

The some 7,000 incidents which teachers provided have been classified into ten content categories. The content categories and the percentages of incidents recorded under each are given in Table 1. Let's look at the first column of percentages there.

If we assume that what teachers choose to list is somehow related

to what they value, we can conclude that these first grade teachers value the intellectual over the nonintellectual, verbal skills over quantitative skills, and personal emotional development over peer relationships. We note too that conformity to classroom procedures (our category closest to old-time "conduct") is still highly valued by teachers — or at least problems concerned with "lining up," "not talking out of turn," and so on are cited more often than problems

TABLE 1

"READINESS" INCIDENTS CITED BY FIRST GRADE TEACHERS
(No. of incidents = 6708)

	Percentage of All Incidents		Percentage of Incidents	
			Positive	Negative
INTELLECTUAL CRITERIA		59	42	17
Verbal skills and understandings	24		19	6
Quantitative skills and understands	2		2	<1
Graphic skills	3		2	<1
Performing arts skills	<1		<1	<1
General intellectual functioning	29		19	9
NONINTELLECTUAL CRITERIA		40	18	22
Attitudes toward school and school work	11		10	2
Conformity to classroom procedures	9		3	5
Personal emotional development	14		3	11
Peer relatonships	2		1	<1
Motor coordination and physical condition	4		<1	4
NOT CLASSIFIED		<1		

in motor coordination or quantitative or graphic skills. ("Graphic skills" by the way includes painting, drawing, making letters–the execution of lines or marks on a surface.)

The last two columns of Table 1 show the percentage of positive and negative incidents in each category. As would be expected, there were more positive than negative incidents cited, simply because the teachers marked more children ready than not ready, and the rules of the game specified that they were to substantiate their ratings by incidents. (The actual ratio was 3 to 2.) What was not quite expected was that teachers gave proportionately many more positive incidents in the intellectual than affective areas. In other words, *good* social adjustment and *poor* cognitive development were less productive of teacher incidents than their converses. This is a puzzle and much more exploration of the data is needed. (Inci-

dentally, these data are very hot off the computer and almost as strange for me to contemplate as for you.)

The frequencies with which incidents were cited are being further analyzed in terms of characteristics of the pupils who were rated, and characteristics of the locations in which the ratings took place. In other words, we are entertaining the hypothesis that the criteria of school readiness may be different for teachers with different amounts of training and experience, in cities of different sizes, and for boys versus girls, younger children versus older children, children from different socio-economic backgrounds and amounts of pre-school experience, and so on.

Our second detection device, the group discussion, was one we are more accustomed to using, especially as we define test specifications (which are very close to educational objectives and criteria, we hope). A group of child development experts[7] met in Princeton to tackle the problem: What should be the characteristics of a comprehensive set of instruments designed to assess children's readiness to enter usual school programs? They were asked to try to specify the categories of behavior which should be included and to attempt to weight the categories in terms of importance.

The committee, under the leadership of Dr. Miriam Goldberg of Teachers College, took an interesting tack in their deliberations: They tried to outline the kinds of tasks and situations with which children would be confronted during their first few days in the first grade and then relate these to the kinds of abilities, skills, etc., children would have to have to cope successfully with the tasks.

You will be interested to know that the committee was unwilling to ascribe anything but equal weights to characteristics in the three major categories: sensory/motor, cognitive/intellectual, and social/personal. They felt that measurement of physical characteristics should best be left to appropriately trained medical personnel, but that criterion instruments could be developed which teachers could use to assess 9 cognitive/intellectual areas, 10 social/personal areas and 4 areas which cut across two or more of the other main areas. The major categories which they produced are given in Table 2.

Of course, to qualify as useful criteria of "school readiness," each of these categories has to be spelled out in more specific terms. An example of a first stage in such spelling out is given in Table 2 for the crucial cognitive category, *Translation from one code to another.* The next stage will involve identification of what letters and speech sounds are to be discriminated, and so on.

TABLE 2

PROPOSED DOMAIN OF AN ASSESSMENT OF CHILDREN ENTERING FORMAL SCHOOLING

MOTIVATION, ATTENTION*

Sensory/Motor**	Cognitive/Intellectual	Social/Personal
Motor behavior	Language comprehension	Impulse control (adaptability, ability to withhold or omit a response)
Functional aspects of vision, hearing	Language output — repetition, restructuring	Tolerance of ambiguity
Nutritional disorders	Concept development and formation (including classification behavior)	Attending to expected task
Etc.	Translation from one code to another	Conforming to and remembering rules and routines
Outlined further as follows:	Ability to absorb, store, retrieve, and utilize information	Fitting into the group and the pecking order
Visual discrimination of letters—scanning strategy, recognition of nonreversibility of letters	Loss of egocentrism (in the sense of recognizing objective dimensions of the world and understanding that other people have other points of view)	Understanding the reward and punishment system to which they are subjected
Auditory discrimination of speech sounds	Evaluation of performance against a standard — performance of self and of others; external and internal standards	Giving up egocentrism (in the personal-social sense; see preceding column for comment on egocentrism in a primarily intellectual sense)
Recognition of regularities	Problem solving strategies (multiple functions, including inductive and deductive strategies)	Sex role behavior
Ability to induce regularities	Ability to initiate structure	Persistence — "give up" rates in the face of failure and success
Ability to identify reusable elements		Ability to role play (an aspect of self-concept)
Ability to name letters		
Visual recognition of words (including recognizing things that are words)		
Ability to recognize speech and language units		
Ability to identify sounds associated with letters		
Ability to attend to auditory and visual stimuli (filter out the static or noise)		
Also identified as important:		
Understanding the purpose of reading		
Picture literacy		

Following directions

Shifting set

* Related to or pervading the other three areas.
** To be checked by appropriately trained medical personnel.

It has been of interest to compare the incidents cited by first grade teachers with the categories developed by the assessment committee. The overlap is considerable, with the teachers providing some rather clear examples of what the committee had in mind, and thus direct help in spelling out the committee's recommendations in more specific terms.

For example, a teacher's "Expresses ideas in complete sentences" conveys some of the sense of the committee's *Language output.*

Negative and positive instances of *Concept development and formation* are provided by these two teacher statements:

> In a series of four houses, she could not locate one that was different. (The door and window were in reverse positions.)
>
> When we cut a pumpkin for a jack-o-lantern, he told us to save the seeds to plant next spring so that we could have a pumpkin for next Halloween.

Positive examples of the category *Translation from one code to another* are provided by these three teacher instances:

> Can interpret the feelings of picture story characters by observing their facial expressions.
>
> Counted his lunch money correctly the first day of school.
>
> Delighted when he was able to recognize some of his reading words in books at home; said he especially liked the word *look.*

Negative instances of the same category include these:

> Had difficulty distinguishing one letter name from another.
>
> Writes her name with all capital letters, most of the time backwards.

The child about whom the teacher wrote "Unable to recall what comes next in a story that has been read several times" may have difficulty with *Language comprehension* and *Ability to absorb, store, and retrieve information.*

On the social/personal side, we believe that the child who "Went home one day at recess, thought school was out" was having some trouble *conforming to and remembering school routines.* "Waits for her turn" seems to say something about *Fitting into the group and the pecking order,* and the child who "Hums to himself continually" seems to be lacking in *Impulse control.*

So far we have involved several hundred people and many thousand dollars in this attempt to pin down "school readiness," and, as

you can see, we are still not done. We are coping first with the "magic of words" problem. What eventually comes out of this work at what is more or less a national level, will have to be considered in terms of its relevance for local school systems and specific educational programs. And there are still the measurement problems to be solved. For example, we are not content to say that *Tolerance of ambiguity* is either too intangible to measure or that satisfactory ways of measuring it have been developed.

The present project simply serves to illustrate what a tough, time consuming job it is to specify the things we want to accomplish with our educational programs, and to define the criteria by which we shall know whether we've accomplished them. The alternative, of course, is to continue going along, hoping we will get somewhere if we only walk long enough. If we take the alternative, however, how are we going to hold up our heads and preserve any semblance of professional integrity as we are bombarded with the sensible questions asked by a sensible public? "Where is *somewhere* and how long is *long enough*?—with Head Start, with new elementary and secondary curricula, with the ungraded school, with the language laboratory, with computer assisted instruction, and with all the other educational innovations which are clamoring for enormous amounts of support.

In case you think it has been a little frivolous of me to lean so heavily on Lewis Carroll in this speech before a distinguished educational audience, let me read you five quotations:

1. We had the best of educations—in fact we went to school every day . . . with extras . . . French, music, *and washing*—extra.
2. . . . Excellent at skipping but deficient in resting . . .
3. Robert needs to feel more sure of himself before he can settle down to the more academic aspects of kindergarten . . .
4. Reeling and Writhing, of course . . . and then the different branches of Arithmetic—Ambition, Distraction, Uglification, and Derision.
5. . . . Deficient at cutting and pasting . . .[8]

Two of these quotations are from *Alice in Wonderland,* the other three from kindergarten teachers' reports to parents. Now I ask you which is from which?

REFERENCES

1. Carroll, Lewis. *Alice's Adventures in Wonderland.* New York: The Modern Library (Random House). P. 86.
2. Dyer, Henry S. "The Discovery and Development of Educational Goals," in *Proceedings of the 1966 Invitational Conference on Testing Problems.* Princeton: Educational Testing Service, 1967. Pp. 12–24.
3. Fuller, Elizabeth M. *What Research Says to the Teacher: About the Kindergarten,* No. 22. Washington, D. C. National Education Association, 1961. Pp. 7, 16.
4. Mager, R. F. *Preparing Objectives for Programmed Instruction.* San Francisco: Fearon Publishers, 1961. P. 44.
5. This project is being carried out at Educational Testing Service by Scarvia Anderson, Ann Jungeblut, George Temp, and Donald Trismen. A complete report will be published later this year.
6. For example: Kohlberg, Lawrence. "Early Education: A Cognitive-Developmental View." Unpublished multigraphed manuscript, 1967.
7. Miriam L. Goldberg, Columbia University; Allan Hartman, Eastern Regional Institute for Education; William O. Jenkins, The City University of New York; Gertrude Justison, Howard University; Harry Levin, Cornell University; Henry J. Mark, Johns Hopkins University; Irving Sigel, Merrill-Palmer Institute; Martin Spickler, U. S. Department of Health, Education and Welfare; Sheldon H. White, Harvard University; Ralph Witherspoon, Florida State University.
8. Quotations 1 and 4: Carroll, Lewis. Op. cit. Pp. 121–122. Quotations 2, 3, and 5: Teachers' reports, *New York Times,* Section C, p. 34, January 11, 1967.

Are Educational Researchers Really Prepared to Evaluate Educational Programs?

RODNEY W. SKAGER

I am facing the problem sometimes encountered by a speaker after following a number of other speakers because there has been some continuity in what has been talked about today with respect to evaluation and innovation. I find that on at least two points something I am going to say has already been anticipated, indeed almost said in exactly the same way. I want to assure you that I have not been sitting back and taking notes so as to be able to jump on the bandwagon.

This afternoon I find myself in the somewhat uncomfortable position of having to fulfill an obligation to speak at length on something I do not know very much about. While this situation may not be entirely unprecedented in the history of educational meetings one does not like to contribute unduly toward turning this particular precedent into a tradition. Perhaps the one thing that can be said in favor of sticking one's neck out in this manner is provided by the experience of those of us who have tried to live up to the role of the teacher. By being continually forced to make noises with our mouths before rooms full of people, we usually work harder to make those noises meaningful than we might have done left to our own devices. I only hope that this general principle will be operating this afternoon.

During the past year I have found that many of the ideas and habits associated with my work have had to be re-examined in the

light of a new situation. While I suppose this happens to most of us from time to time, the particular reorientations I am being forced to make may turn out to be a relatively common experience in the future. Indeed, if this is not the case, I expect that the educational research community will fail to meet a real public responsibility and will also miss a significant opportunity to expand basic knowledge.

As everyone knows, massive infusions of funds earmarked for innovation and development are being administered to American education. The biggest single factor in all of this—though by no means the only one—has, of course, been the Elementary and Secondary Education Act of 1965. It is probably not entirely a coincidence that national support for the try-out of new-fangled ideas and gadgets comes at a time when American industry has begun to get into the business of educational technology in a very big way. So there are plenty of things for the educator to buy and since many of the items are expensive, the matter of evaluating their effectiveness becomes more important than ever. But even if educators themselves were not concerned with evaluation, the Federal Government is concerned. Note, for example, the very significant clause under Title I of the 1965 Bill which requires:

> That effective procedures, including provision for appropriate objective measurements of educational achievement, will be adopted for evaluating at least annually the effectiveness of the programs in meeting the special educational needs of educationally deprived children.

I am sorry about that prose. Next time I will try Lewis Carroll.

This clause, and others like it, have already spawned hundreds and will eventually spawn thousands of evaluation projects. At the Center for the Study of Evaluation of Instructional Programs we have already received many inquiries as to whether we would be interested in conducting the evaluation phase of a local project. Although evaluation needs have been with American education for some time, I think it is safe to say that, within the past five years, they have developed the potential for requiring far more in the way of professional resources than the research community has available now or in the forseeable future.

What I am going to discuss in the next few minutes amounts to an elaboration of the assertion that educational researchers as a group are not necessarily well equipped for the evaluation of in-

structional methods and practices. In case you are anticipating certain remarks, let me hasten to add that I am not going to bore you with another public lamentation about how our graduate schools do not offer or require enough background in research methodology and philosophy of science. This is most certainly true. However, excuses to the effect that students in curriculum administration or media do not need much in the way of research training are sheer nonsense. *Everybody* is in the research act nowadays. Once a student gets a degree and a professional title he is not likely to be challenged in the laboratory doorway by a guard whose assignment is to ascertain that everyone seeking admission had Statistics 210C. So, even though I could make vigorous noises on this particular topic—and I have almost gotten excited—it is really another issue. Statistics 210C may be a necessary prerequisite for research in evaluation but it is not a sufficient one. In fact, it is precisely some of the habits of thinking developed in Statistics 210C and Measurement 211C that are part of the problem. The thesis of this presentation is not that there are too few adequately trained researchers but rather that the well-trained, technically competent researcher may bring habits and values to evaluation research that are out of mesh with the problem at hand.

But before I begin talking about the habits, both good and bad, of educational researchers I had better sketch in what I think evaluation is all about. I have been more careful to do this sort of thing since the day I sat down in a certain barbershop in Princeton, New Jersey. When I asked for the usual crewcut the barber replied, "Well, now, let's define our terms." I knew I was in a new kind of town.

The type of evaluation under discussion here refers in the most general sense, as a very broad definition, to the comparison of educational practices and their outcomes against standards and/or against other educational practices. Embedded in this very loose definition are at least two of the factors making evaluation research a different and rather obstreperous sort of "animal." First, there is the use of the term "standards," implying that values have a very significant role in what is studied. Now, I do not maintain that the selection of a problem in so-called basic research does not reflect the experimenter's beliefs about what is worthy of study, but evaluation research obviously implies a concern with values that is raised to the 10th power. It is quickly apparent, without elaborating on

the issue, that the matter of what such a complex phenomenon as instruction in the schools is supposed to accomplish poses the evaluation researcher with imposing problems of selection and emphasis. In contrast to his colleague in the laboratory, the evaluation researcher will find it almost normal for people to have violent feelings about what he did or did not include as a criterion variable.

Secondly, the use of the terms "educational practices and their outcomes" indicates a most significant fact. *What is done* is just as much the subject of evaluation research as *what happened when it was done.* In reports on laboratory research, the description of the experimental treatment is never the subject of the study; it is merely part of the introduction to the report in the journal. But in evaluation research it is at least half of the battle. It does no good, you see, to say that school A produced more knowledgeable students than school B if you cannot tell people what it was that the two schools did that was different. It is pointless to report that curriculum C produced better results than curriculum D if you cannot establish that both curricula were actually presented in the way you think they were presented. We all know that human beings can very easily give lip service to new procedures yet really do things in the same old way.

I think one more thing needs to be said about the definition. The terms are deliberately as broad as possible, reflecting the nature of the potential subject matter. I have, for example, had to use the term "outcome" instead of the term "learning," not only because our criteria about what happens to students may encompass traits other than those usually characterized as "knowledge" or "skill," but because much of our data on outcome may not involve the student at all, either as individuals or in the aggregate. We have to be concerned with such diverse outcomes as changes in the structure of the school as an institution, the behavior of teachers, or even—perhaps I should say especially—whether a new program costs more than it is worth.

In recent years, some general notions about the particular characteristics of evaluation research have arisen. As far as I can tell—and I am a neophyte in this field—Cronbach's (1963) paper on curriculum evaluation[1] has been the fulcrum around which these ideas have developed. Cronbach advanced, among other things, the radical notion that the traditional emphasis on evaluating final outcomes has been all wrong—that is, Cronbach rejected the image of the evalu-

ation researcher as one who sits back until the developer turns out a finished product and then, entering the picture at its most dramatic moment, designs a critical experimental test, involving, to be sure, the random assignment of subject, and from his Olympian heights, eventually dispenses the masterful conclusion that group E performed better than group C at the .05 level of confidence. Cronbach suggested that image was very nice, but is seriously flawed in practice by the fact that the usual result is the retention of the null hypothesis rather than its rejection. On the rare occasions when the hypothesis of no difference is rejected, one has the feeling that some minor tinkering around with the curriculum or other method which came out most poorly might have changed the result radically.

Cronbach's answer to what the evaluator ought to be doing is well known and shared to varying degrees by most people who are presently writing about the matter. The evaluator should be concerned primarily with using evaluative information in a formative sense, that is, he should enter the picture early in the developmental phase, not at the end. He should observe how things are taking shape, provide information that will illuminate decisions about adopting particular strategies or methods, and, in general, sharpen up the developmental process. The idea is that the evaluator's primary role is to help see to it that the product eventually turned out is maximally effective, not just to aid in making a decision about adopting that product. This is a dynamic and constructive role, as opposed to the, at best, disinterested involvement of one who judges the final product against other alternatives. Perhaps I might anticipate myself at this point by suggesting that the hypothesis testing psyche of the typical researcher—which implies an orientation to issues and questions rather than materials and services—is hardly attuned to the type of role Cronbach has envisioned. Although others, particularly Scriven,[2] (1965) have been unwilling to throw out evaluation of final outcomes—and I suppose I would count myself on that side of the fence—everyone these days sees this constructive or formative role as at least a major function of the evaluator.

There is another emphasis in contemporary theory about evaluation that will help us understand an aspect of the beast that makes many experimenters uncomfortable. And this point of view was well articulated by Guba[3] (1965) in his paper on strategies for producing educational change. In this paper, Guba strongly emphasized the

idea that evaluation research is properly done outside the laboratory and in the natural setting in which the program was designed to operate. Rather than the precise control over extraneous factors available in the experimental situation, evaluation research should, in effect, toss down the gauntlet and dare the context to confound the results. After all, the context will always be there when a program is adopted in the schools. You might as well study it. Referring to Barker's[4] interesting distinction between the interventionist role of the experimenter and the noninterventionist role of the observer–transducer in Barker's terms–Guba rejected the experiment as the primary vehicle for evaluation research and thus, in effect, said that the researcher in evaluation should be especially careful to avoid altering the phenomena that he is observing by setting up the artificial constrictions that an experiment implies.

In reply to this view, I think it should be pointed out that even if we are entirely convinced that the effectiveness of educational practices can be determined only in the natural setting, we have by no means eliminated the experiment, only the laboratory. It is certainly possible to perform experiments without creating artificiality. Control is, of course, a greater problem in the so-called real world, but control approaching that of the true experiment, and certainly the quasi-experiment, does not inevitably require the researcher to change the conditions under which the instruction takes place. Further, if we allow the simulation study as an evaluation tool, I am not even certain that we have eliminated the laboratory. After all, if the evaluation researcher is to aid in the formative aspects of the development of educational programs, it is hard to see how we can eliminate techniques that are essentially simulation as opposed to "real" instruction. If we vary conditions and compare results in simulation research, we are back to the experiment in a situation that is to some extent artificial.

While I am by no means willing to agree that the experiment has no role in evaluation, I do believe that the otherwise competent researcher will put far too much faith in its value in his initial approach to an evaluation problem, and, even where experimentation is appropriate, go about things in a way that does not necessarily meet the requirements of a new situation. The first difficulty in utilizing the experimental method in evaluation has already been mentioned. How valuable, really, are experiments in which you are uncertain as to the nature of the independent variables? Comparing

the learning of students under two different experimental curricula may, even in the face of Cronbach's doubts, show that one program is superior. But each program is a complex entity, subsuming many subsidiary variables. For the purposes of the experiment, we can call the entire program a variable. However, when we do this, we are in danger of committing the error made by the man who wanted to relegate the automobile to some sort of primeval junk heap the first time one lost a race to the horse. All that was needed was a better spark plug. Instead, the whole vehicle was condemned. The results of an experiment pack tremendous power for establishing conviction, but we had better be certain that we know what it is we have been experimenting with.

In the typical laboratory experiment, the researcher can determine in advance everything that is going to happen, except, presumably, his results. (And there are some who might say that he often does that as well.) In short, conditions can be standardized in the laboratory, contingencies are taken into account, and the whole procedure can be so automatic that the experimenter attends a convention while his graduate assistant runs the experiment. Perhaps I am being facetious, but it is actually true that the well-run laboratory experiment ought to go off like clockwork. Both independent and dependent variables are determined in advance and are usually relatively few in number. The subjects may be doing many things during the experiment, but only a limited set of variables are recorded and studied.

Now, I submit to you that the experimenter who generalizes these habits to the nonlaboratory evaluation situation is almost certain to produce a very uninformative, perhaps completely misleading, evaluation study. The non-involvement of the experimenter with the subject matter of his study—as might be typified by the anonymous rat in the programmed Skinner Box—is simply an inappropriate model for most evaluation research. Let me hasten to say again that I am *not* saying that the experiment, even in the purest sense of the word, is always inappropriate for the evaluation of instructional programs. But I am saying that some of the habits and expectations of people who do experiments (and train others to do experiments) may not be appropriate in evaluation research.

The researcher must, therefore, resist the tendency to treat complex instructional totalities such as a curriculum or a school as if they were the simple, unidimensional independent variables usually

dealt with in the experiment. Furthermore, he should realize that the description of just what the major instructional variables are may well be the most important single outcome of his research. And finally, the researcher should avoid the premature crystallization of decisions about what is and what is not to be observed and substitute, at least in part, an active and flexible involvement with the educational situation under study so that measures of variables, both dependent and independent, can be collected, even though their importance was not anticipated at the beginning of the research.

Needless to say, this tradition of the highly standardized data collection phase is not limited to the experimentalist. An even greater offender is the survey researcher who designs and mails out a questionnaire from the splendid isolation of his office. His closest involvement with the subject matter may be a casual flip through several pounds of computer print-out. Let me make it clear that there is nothing wrong with this as long as we are all-wise in advance. But I would maintain, however, that the complexities of educational treatments and their many-faceted effects make it very difficult to take everything into account in advance.

But habits associated with the laboratory experiment are not the only problem. Everyone in educational research, even including the lofty psychometrician, is likely to have some habits or beliefs that do not integrate well with the requirements of evaluation.

As an opener, I would like to suggest that there is just a faint possibility that we have oversold what are usually referred to as "behavioral goals," and here I must add the prayer that heaven may not strike me dead on the spot. By "we" I refer to the research community, and by "behavioral goals" I mean something like the idea I heard expressed recently by a man representing an important governmental agency. "If an instructional program doesn't make any difference in the learning of students," he said, "we should throw it out and spend our time looking for something better." While a year ago I would have agreed with this assertion without giving the matter a second thought, I now feel the uncritical acceptance of this statement by evaluation researchers is erroneous. For one thing, the focus on the word "learning" needs to be elaborated. I think the word "behavior" in "behavioral goals" should have a broader referent.

Take the ungraded school, for example. One of the things that

this sort of organization is supposed to generate is a greater motivation toward independent study on the part of the student. To be sure, learning is a goal, to some extent an ultimate goal, but would we not be interested in getting some measures of the manner and frequency of independent study activities—and I do not refer to anything elaborate here—as a perfectly appropriate criterion measure? Moreover, I would be willing to let the measure stand by itself as being at least equal in importance to a measure of what has been learned. After all, independent study may tend to personalize learning to some extent. Give a typical broad gauge achievement test to a classroom full of independent learners and you may miss the mark in many cases. Unless you are insightful in designing the learning measure (for example, establish the dangerous precedent of asking the student what he studied before making up the test), you may be unable to confirm the result you had expected.

But there is another aspect of this overly restrictive definition of "behavioral goals" about which it may be harder to generate conviction among researchers. I have anticipated this point already by emphasizing that evaluation criteria need not be restricted entirely to measures taken on the student. This proposition is advanced with some misgivings, because I certainly do not want to give the impression that the traditional approach to evaluation based on long check lists about the facilities and procedures, without any reference as to how people are behaving, has much validity. But there are many things going on in the school that are highly significant, that can be assigned values, but which are only tenuously reflected, if at all, in the learning of students. For example, we would not have to give achievement tests to students to decide whether it is a good thing that a 20 percent increase in annual rate of teacher resignations accompanies the ungraded school. To be sure, we will want to check on whether we have instituted a wholesome process of natural selection accompanied by "the survival of the fittest" of our teachers. But measuring our students' learning seems to me to be a very indirect way of determining this. There are many goals that ought to be studied primarily at the institutional level. For example, determining the effectiveness of the League of Cooperating Schools, set up by the Institute for the Development of Educational Activities, involves legitimate criteria that are rather far removed from the behavior of the student. As an organization set up to stimulate innovation and change in the schools, the

League's efforts must be evaluated, at least in part, simply on whether any changes have actually occurred. For example, if it can be shown that teachers are meeting regularly to discuss curriculum innovations and if this seldom or never happened in the past, then I think most of us would feel that at least one positive outcome had been demonstrated, even in the face of the fact that students in the school still have the same mean score on the STEP test. Being broadminded about our criteria in evaluation does not imply a lack of intellectual or moral rigor. It merely indicates that we have recovered from an inhibiting case of tunnel vision.

With respect to criterion measures, narrow-mindedness about the range of variables relevant to evaluation is only one of the potential bad habits of the typical researcher. Considering now, the matter of student learning, I would suggest that there is a tendency to rely too heavily on standardized psychometric instruments designed for purposes other than the evaluation of instructional programs. I am, of course, referring to the achievement test. To be sure, we rely on the achievement test partly because measures specifically appropriate to many criteria do not exist. In the effort to get something measured, the evaluation researcher may select his criterion instruments largely on the basis of what is available. In effect, the state of the measurement technology establishes the criteria rather than the investigator.

But the state of the art is not the problem referred to here. Rather, the difficulty is that most of us have certain fixed, though largely inaccurate, notions about the properties of the achievement test. In particular, we have the idea that such instruments are designed to be sensitive to what goes on in the classroom. This is untrue. The point has already been made by many individuals that the manner in which achievement tests are built actually works against this goal, not intentionally, you understand, but as a natural result. The achievement test is useful for the purpose for which it was designed: ordering or discriminating among individuals with respect to knowledge and skill associated with rather broadly defined areas of educational content. Oddly enough, evaluating people, which is what the achievement test was designed to do, does not turn out to be the same thing as evaluating instructional programs. We would heap scorn on an achievement test if it turned out that most of our students got almost all of the items right. Our legitimate complaint would be that the test had failed to discriminate among

individuals. For this reason, as well as to promote test reliability as traditionally defined, test makers tend to select items likely to be passed by about half of the subjects to whom the test will eventually be administered. This is fine for the discrimination application of achievement tests. But it is a perfectly silly way to build an instrument for determining whether we have accomplished certain instructional goals. In the latter type of evaluation, we would, presumably, be quite interested in a test on which the majority of our students had nearly perfect scores—assuming that our efforts at instruction were successful. In fact, long before I took measurement courses there used to be a term for such instruments. I believe they were called "mastery" tests.

But the matter of test difficulty is only one problem associated with the use of achievement tests in evaluating instructional variables. The bald fact of the matter is that scores on achievement tests are just generally insensitive to particular educational experiences. Achievement tests tend to reflect rather stable characteristics of the *learner*. This is one of the reasons such tests are reliable, and it constitutes an absolutely essential condition for the evaluation of individuals. Nevertheless, there appears to be a rather pervasive belief that intelligence tests and achievement tests are very different sorts of measures, that achievement tests reflect school learning, while intelligence tests do not. Actually both types of instruments reflect school learning, but mainly in the sense of being related to more general variables such as number of years of schooling in the case of both types of measures, or whether the student has taken a given subject in the case of achievement tests. After all, the high correlations between aptitude tests and achievement tests reported in test manuals and many research studies ought to convince us that the two types of tests reflect many of the same aspects of students, both cognitive and motivational. All of this adds up to the fact that the use of total scores on achievement test in the evaluation of instructional programs often may insure negative results, even when the program might be operating quite effectively.

You might have the idea by now that I am saying that achievement tests should not be used in evaluation studies. This is not true at all. I think that standardized achievement tests ought to have extensive application to the evaluation of instructional programs. What I *am* objecting to is the uncritical use of total scores on achievement tests in the belief that such scores will be sufficiently

sensitive to what goes on in the classroom to provide a basis for deciding on the effectiveness of particular educational practices. If we forget about total scores, however, and look at the level of the individual test item, it is quite possible that the achievement test will turn out to be a very useful tool in evaluation. Why? Well, although total scores on a test or scale may be little affected by the program we are studying, a few of the items may show dramatic changes in proportion passing or in correlation with some relevant criterion. Presumably, these items would be more closely related content-wise to the program under study than is the test as a whole. At the purely empirical level, then, analysis of individual items would provide us with a means of discovering what particular learning has occurred. Such specific information is far more meaningful and relevant to the task of evaluation as defined here than in any total score that I know of. If we combine this search for changes on individual items with what are now referred to as "item-sampling" procedures, the latter allowing us to administer very large numbers of items, but each item to only a small number of individuals, we then would have an evaluation tool of really interesting potential. I would admit that the tool looks suspiciously like a shotgun, but at least it is a vastly bigger and more powerful shotgun.

While I may have been somewhat hard on the researcher up to this point, I think we all recognize that there is another side to the coin. The hardheaded research type, if we can induce him to involve himself in the evaluation of what we do in our schools, may have some points of view, in addition to his technical competence, that will make both a unique and significant contribution. I would like to speak about two characteristics which the researcher, perhaps more than anyone else, has to offer.

First of all, the researcher who is worth his salt is likely to be interested in at least some aspects of theory in the behavioral sciences. If such interests influence his approach to the evaluation of instruction, and they should, the researcher is likely to pose some fundamental questions about what the program is accomplishing—questions that otherwise might not have been asked. For example, the developmental psychologist may be interested in finding out whether a given program can teach first graders to perform certain arithmetical operations—straightforward instructional goals—but he may also become curious as to whether the instruction has any

influence on cognitive processes at a level more basic than instructional content. For example, Piaget, whether or not you accept his theories about cognitive growth, sees the development of certain conceptions as naturally antecedent to learning the formalized operations of arithmetic as well as other rather different, educational content. One would expect that a researcher interested in these more fundamental concepts and processes would manage to find ways to study them as a part of his efforts at evaluation. On the one hand, he might collect evidence confirming Piaget's notions that children who have not developed certain fundamental skills cannot learn arithmetic. This information would be useful to the developer, and also would make a contribution to knowledge in what is usually referred to as basic research. On the other hand, the researcher might discover that the program influences the development of the fundamental cognitive processes he is studying. Such a finding would obviously be of even more importance to all concerned.

The point of all this is that research, *any research*, involves asking questions. The researcher is a man who, if he is any good, asks the most interesting questions—even some questions that, at first sight, might not appear to be relevant to the practical problem of making a program work. At best, he will relate the evaluation of instructional procedures or techniques to issues that transcend immediate practical concerns.

A second contribution that is in the special province of the researcher—and again this is in addition to technical know-how—also involves the questions asked in evaluation. Scriven[2] (1965) has pointed out very forcefully that there is a widespread tendency to "relativize" the evaluation function—that is, there is often pressure to restrict evaluative information to criteria generated primarily by the particular educational program under study, rather than by instructional goals with broad generality. This, to be sure, is an impressionistic observation. It cannot be verified by citing a U. S. Office survey of the goals of evaluation projects over the past five years. But I think that anyone who has had much contact with developers will have run into this tendency to view many criteria as either biased or irrelevant to the program under study.

Developers are naturally enthusiastic about their own ideas and sometimes rather scornful about what has been done in the past. It is especially significant that not only are earlier instructional *methods*

viewed as inadequate, but earlier *goals* as well. It is often possible that the developer's enthusiasms will limit the thrust and scope of the questions posed in the evaluation.

I view this tendency to "relativize" evaluation goals as, first of all, unrealistic from the pragmatic point of view. It is obvious that concentration on a limited set of instructional goals in evaluation is likely to limit seriously the spin-off of discoveries in the behavioral sciences. But sticking with pragmatics, we ought to recognize that people making decisions about adoption of the program in schools may have questions about other goals. We may accept the instructional intent of a modern mathematics curriculum, but still wonder how students taking the new curriculum fare in terms of the goals of "traditional" mathematics instruction, whatever they may be. To agree that the traditional criterion may be biased against students who have taken the new curriculum is emphatically not a reason for failing to include that criterion in our evaluation. The presumption of bias requires us to be careful and responsible in our interpretations, but does not indicate that the information provided is without meaning.

It is my belief that concentration on a program-specific set of goals will run against the grain of the average educational researcher. If he is at all theoretically oriented, he should be accustomed to thinking in terms that are considerably broader than particular instructional content. Thus, the most important strength the researcher brings to his role in evaluation is not his knowledge of psychometrics, or statistics, or experimental design, but the simple fact that he is by nature a creature who asks questions. In the light of this, it would be helpful if diplomacy were also one of his strengths, because those questions may, at times, make the developer uneasy.

We have seen that the behavioral scientist may have some things to learn before he can perform effectively in his role in the evaluation of instructional programs, it is equally true that he has significant things to offer in addition to his technical competence. But one cannot help but wonder how the evaluation researcher's findings contribute to final decisions about which innovations are to be retained in the schools. It may be that this sudden escalation in our evaluation requirement is creating a new role for the educational researcher, one that is somewhat closer to the final decision process than any he has played previously.

Now, the school is, and always will be, a socio-political institu-

tion. The influence of the researcher on the men and women who make decisions about curriculum, school organization, and the like, is, in one sense, similar to the economist's influence on the government official. A tax boost this year may be excellent economics, but nonsensical politics and, thus, overruled by that official. I am sure that the same situation often pertains to many of the innovations we in educational research see as imperatives. Yet, in another sense, I do not feel that educational researchers in the past have had quite as formalized a role in the educational decision-making process as economists and other experts have had in government. This is not to say that research has not had a genuine impact on educational policy. It most emphatically has. But the influence has perhaps more often conformed to the traditional "pipeline" model in which ideas or inventions are generated in the scholar's armchair or laboratory, go through a developmental phase (usually in the hands of others), and then, as formalized procedures, go through dissemination and adoption phases. Insofar as this model is appropriate, it portrays a rather wide separation between the scholarly researcher and the educational decision maker.

Much of the substance of what I have had to say this afternoon amounts to a contention that the evaluation researcher, if he is to be effective, will have to get a lot closer to the school in all its aspects than he has in the past. It seems to me that this requirement is also going to put many researchers very much closer to the times and places where decisions are made. In other words, the researcher, perhaps in spite of himself, has been washed a considerable distance closer to the other end of the theoretical pipeline.

I have a sociologist friend, who, the other day, explained to me that all of these observations are perfectly obvious to anyone who is fortunate enough to be a sociologist. If I understood him correctly, and there is some question as to whether I did, it works something like this. Roles are defined by people's expectations. In other words, there is some sort of consensus that people who occupy a given role ought to behave in certain ways. Much of this paper deals with what we have a right to expect from the evaluator. Although defined in terms of expectations, roles are created to fulfill functions, that is, to accomplish something. Functions, in turn, stem from demands made by society. You may say, "Well, what demands has American society made that have generated or at least stimulated the evaluation role?" I think the answer is clear. Very powerful

voices in our society are saying, "Do something to adapt the schools to the educationally deprived!" "Do something about incorporating the new technology into instruction" and "Do something about applying new ideas about the learning process." Political leaders have, in part, responded to these demands with the Federal Education Act of 1965. With its provisions for the annual evaluation of educational programs that are already in operation this Act has gone a long way toward creating a new role in education for the behavioral scientists, one that is clearly closer to the places where decisions are made. If we who are in research continue to take some time to think about the nature of this new role, we will be much better equipped to fulfill it.

REFERENCES

1. Cronbach, L. J., "Course Improvement Through Evaluation," *Teachers College Record*, 1963, *64*, 672-683.
2. Scriven, M., *The Methodology of Evaluation*. Multilithed Report, 1965.
3. Guba, E. G., "Methodological Strategies for Educational Change." Paper presented to the Conference on Strategies for Educational Change, Washington, D. C., 1965.
4. Barker, R. G., "Explorations in Ecological Psychology," *American Psychologist*, 1965, *20*, 1-14.

Responsibility in Innovation

LELAND B. NEWCOMER

I feel I am in a rather ridiculous situation. I am not a scholar and I feel just a little bit out of place. I guess if I were to attempt to classify myself I would like to be known as a "manager of educational enterprises" or maybe an "organizer of time, space, people and things" or perhaps a "politician."

There are a lot of people who are superintendents who would not admit that they are politicians and I feel that that is one of our problems in education today. We have superintendents who are not too honest—don't misunderstand me—by that I don't really mean a moral virtue, I mean they tend not to face up to the critical issues of our time and they don't face up to the fact that public education is in the political marketplace. We educators are in the heat of the kitchen. I for one think this is good. We have been fighting for many, many years to get there and have been saying how important education is. It is the most important thing that goes on.

If you include what is called the "knowledge industry"—book publishers, publishers of instructional materials, technology of various kinds, including the expense of public education, you have the biggest industry in the amount of dollars spent in the whole of the United States, even surpassing the defense effort, the war effort and everything else.

With this being true, we are kidding ourselves if we don't face up to the fact that we are going to have to be held accountable. We are in a position where somebody is going to be looking at us all of the time. If we don't choose to be in that kind of dilemma I suspect we should not choose to be in public education or at least not choose to be a superintendent.

I am really happy to have the opportunity to talk here. I relish the opportunity, not because I feel qualified, but because I do have a lot of ideas which, although I'm sure I can't prove and which are not even statistically measurable, are ideas that people ought to think about. I am not even convinced about them myself but I talk about them as though I am very much convinced because that tends to make people react more violently and stirs up their adrenalin a little bit so that more thinking goes on.

I am one of those people just crashing for innovation. I want to innovate all over the place. I think it is essential. I think the way teaching and learning goes on now is pretty obsolete, really. I got into trouble the other day by saying that we were operating at 50 percent efficiency so I won't say that today. But the important thing I want to talk about is "Responsibility in Innovation." I am not going to talk about sophisticated research or sophisticated evaluation because I really don't know anything about that. I am going to attempt to talk about the organization of schools. I don't know much about things other than schools. I would like to talk about an organization of schools which will increase the amount of, and the speed of, change or innovation. I hope that this organization will increase the chances that what we do differently will make a difference—a difference toward achieving the objectives we have.

The other day I was talking to Dr. Harold B. Gores, president of the Educational Facilities Laboratory, a Ford Foundation activity. Dr. Gores is an extremely popular man. His job is to give away many millions of dollars each year to people who can come up with significant ideas which might tend to increase the chances of making more efficient use of the facilities in our school buildings. He said, "You know, we are in a fantastic revolution. More things are going on now — things are 'popping out' all over. There is tremendous pressure. There is a fantastic increase in the amount of funds provided by the Federal Government and this creates many problems. It creates excitement, it creates resources and it creates action."

I like action. Imagine Title I, II, III, IV, V, ESEA and all the rest. Imagine knowing that all that money is there and imagine the problems being created. What do you suppose goes through people's minds? I have a sneaking suspicion that what tends, on too many occasions to go through the local districts' mind is, "How can we get some of that money?" rather than looking at the situation and saying, "What problems do we have for which a sound proposal

can be made." "How can we define the problem in such a way that we can find the best source for its funding?"

Now, if we do this then I suspect that what we are doing is going to really enhance and increase the effectiveness of teaching and learning. If we take the other route "Get the money" then we are going to end up in just doing one thing and that is, proving how incompetent we really are. We have been screaming for money for a long time. Now we are getting money and if we don't produce something with that money it will be the *coup de grâce* for educators as managers of schools.

You have read Ralph Hemling's article "Is the School Superintendent Obsolete?" He builds a pretty good case. I don't agree with him, of course, but irresponsible innovation is the best possible way that I know of to make the educational enterprise or establishment as it now exists, obsolete.

Let me get more into what I want to talk about. A dilemma exists. We have people running in two directions. We have those that are saying "be careful"—these are the reactionary ones. They are thinking of 125 reasons why they shouldn't or can't be doing something. Then we have those people that say, "Don't just stand there. Do something! Do something!" You know, those are the ones that say "Are you doing team teaching?" "Do you have ungraded classrooms?" "You don't?" "Isn't that too bad." "Do you have continuous progress in your schools?" "Do you have modular scheduling?" "Do you have flexible scheduling?' "Do you have independent study time?" "Do you have . . .?" We could go on and on and on. They are becoming clichés. We have people who are just busting out innovating all over the place because they think it is the "George" thing to do and then we have those who are dragging anchor.

Why does this take place? I say it is because of poor management. You see, I am prejudiced. I think good superintendents are essential if we are going to get something done, but their primary task is not to do things. It is to organize an environment in which these things will happen and happen effectively.

I am sure I don't know exactly how to do this but I have some ideas that I want to discuss with you. Before I get into that specifically though, people have talked about the technological revolution and the knowledge explosion and I think they have been referred to today so I don't have to dwell on that. Just to make this one point, I really believe that we are kidding ourselves if we think

the present organization and the present way we approach public education will allow us to continue effectively. As we are presently organized it is impossible to effectively take advantage of the knowledge and the hardware that we already have in our schools.

The reason I say this is because as I look at school districts I have to assume by the way they are organized, that they are organized for the purpose of control. I have never seen an organization that is so well set up to be conservative or reactionary. Local boards of trustees are reactionary. Local boards of trustees have a tendency to think that when they get five phone calls or when 20 people come to a board meeting and express a point of view that this feedback represents the universe. You get action that way, you know. Well, this kind of thing creates a very real problem because most of our schools are organized this way.

I would suggest to you that perhaps schools should be organized at least 50/50—that is 50 percent of the resources going to control (keeping the store, operating the establishment) and 50 percent of the effort or power or prestige or whatever you want to call it, going to "break-through"—going to the problem of improving the operation. Now I again submit that if this doesn't happen we will become obsolete. Somebody else is going to take over the function of education in our society. I firmly believe this although I don't approve.

The real revolution, as I see it, aside from this organization problem, is a revolution away from teaching to learning. That is an over simplification and I am sure some of you people can poke holes in it. I hope you are thinking about it right now. But the emphasis used to be on teaching. When I first started observing classrooms and doing, so called, curriculum supervision, we would go into a classroom and would evaluate or manage by process or by method. I call it "characteristic evaluation." We would note, "Does the teacher smile?" If the teacher smiles a lot then we infer that the teacher must be warm. If the teacher is warm then that ought to be good for kids and so we say that probably if she smiles it increases the chances that kids learn.

Then, if there are centers of interest, if the pictures are straight and the room environment is good—boy, we really worked on room environment—we thought it was good. Don't get me wrong. All these things have some kind of validity but unfortunately that was the only thing we were using. We were really looking at process and method all the way.

What is the net result of an enterprise that evaluates or manages by the methods used? What happens is that it evolves toward management by prescription. That means that you tell teachers what to do and how. The one thing I fear today is that most of the innovations we have going on now are becoming prescription innovations—team teaching, modular scheduling, flexible scheduling—all these. Somebody who knows all is prescribing a new means and then the people who are to implement the means are becoming like the people that implemented the means before. We attract people who just want to follow the leader.

I think John Dewey was great. I think he had some of the greatest ideas that have ever been contributed to education. But some of his disciples, in my opinion, really fouled up these ideas.

A funny thing happened when I began teaching. The rage then was the social studies unit of work. This was an innovation and it was creative and I think it was sound, but you know, before we got through it was more traditional and more crystallized and more rigid in the prescription of method than anything I have ever seen. It started out as an objective but it ended up as a prescription of method. This is the thing I am worried about. And I think the revolution from teaching to learning is when we put our attention to learning then we can stop thinking so much about the means as we develop the ability to measure the results of learning.

We haven't gotten there yet. Here I am getting way out of my field. I am getting into your field so I am on dangerous ground. I don't think we can measure very many things, but I think there is an in-between step between evaluation or management by method, prescription or process and evaluating or managing by results. And that I call "managing by objective." Earlier Mr. Skager mentioned that the true researcher should ask the right questions. Now, if you can organize so that questions are asked in such a way that the objectives become really clear and you can manage by objective, I think you have made a very strong step toward the eventual management by results.

Let me give you a kind of facetious illustration of what I mean in connection with this. I have an objective. I have a problem. My problem is that I hate Board meetings. I hate Board meetings with a passion. I should probably never have been a superintendent because that is a dilemma for me. Having Board meetings is one of the functions which superintendents have to be involved in. But I hate them and that constitutes a problem for me.

How about a plan or an objective—first an objective. An objective might be to somehow, instead of having Board meetings about every week or two (with a special one in between) to have them once every three months or even every six months as the big corporations, such as General Electric do. That would really be something. So that becomes an objective.

Now a plan. You look at the corporation and say, "How can they do that?" I will tell you how they can do that. In the first place they can do it because they go in every six months or every year for their annual meeting or whatever it is and present an evaluation of the performance of the organization in terms of a profit and loss statement. If they make a profit, great. That is really their objective.

We in school have a problem. We aren't a profit making operation and we haven't learned how to evaluate our product in terms that the community will accept as evidence that we are achieving the objectives they desire. Now, we should be able to do that. I have a sneaking suspicion that I could get Board meetings down to once every three months or once every six months because I could say to them, "Now, look. Let's face it. This we have clearly defined is what we are here for. We agree. O.K.? O.K." We go the next step. We say, "Now, this is what you will accept as evidence that we are doing what you think we should do—that the product is as you think it should be? This is the evidence." Then, I produce the evidence.

And so, every three months we look at it and we see whether it is going in the right direction, whether it is good or bad. If it is bad the Board says, "Buddy, you have three months to get it started the other way." If it is good you might ask for a raise.

Every three months, then the Board members and the community will get their minds off the prescription method. They won't be too concerned about what methods are being used as long as kids are learning effectively.

This is why I am so excited about the National Assessment Program. I am for it and what they are trying to do. It is going to open a can of worms all right and we are going to have a few problems but I have learned one thing since I have been trying to be a manager and that is that one of the best ways to get something accomplished is to create a problem. You know, create a big fat problem. In fact I told the Board members at Newport-Mesa, "If you want somebody to come in and run a smooth ship, don't hire me because I plan to create more problems than I solve." I think this is one of

the functions of leadership. I think that national assessment is going to create some problems but for heaven's sake, what kind of an assessment do we have now? Twelve people come to a Board meeting and the Board thinks that they are the universe and so we change. Is that the kind of input we want to determine, the direction and the goals of our schools?

Let's get back to the specific organization. I think that the large central office, the big research department, the grandiose assessment, the large innovations, the large program change or improvement, district-wide operations are passé. You know, it used to be that you could set up a committee in a large school district and then spend a year defining the objectives and then four years later you would implement a new reading program, a new literature program, a new this-program or that-program and then every four or five years, you would take on a different one. That's ridiculous. Text books are changing faster than that. The innovation has to be at the grassroots. How to get responsible innovation at the grassroots of an operation becomes the real question.

I will tell you how I think it ought to be done. Let's just take a typical organization of the school district. I think if a school district is large enough to have an assistant superintendent they ought to have two. One of them ought to be in charge of operations, an assistant superintendent in charge of operations. He is like the executive officer aboard a ship. He runs the everyday operations. He leaves the captain of the ship free to plan and determine the course and the variety of other kinds of things.

The other one would be the assistant superintendent in charge of "break-through," but let's be a little more traditional and call him "research and development." But "break-through" is why he is there. He ought to be right alongside the other assistant superintendent and he ought to be paid just as much and he ought to be just as much in the line, not staked out in a staff operation some place, but right there in the main stream. His function is to foster breakthrough and responsible innovation.

Now let's take it out of the district office. Let's go out into the high school. By the way, I could talk the whole time on the high school if you would like. I think that is the most beautiful example of Parkinson's Law I have ever seen outside of PTA.

Look at the organization of the high school and the way they have added positions to take care of things that happen without

ever looking back. Let me give you an example. The high schools where I am right now have a principal, an assistant principal and two vice principals. They have only had an assistant principal for about four years. Do you know how they got that assistant principal? They decided they were going to do something to upgrade teaching. They wanted to pay good teachers more than poor teachers and they weren't able to let poor teachers go so they had a problem working this out. They said, "We have to have another staff member," so they hired an assistant principal. But you know, as it worked out they didn't establish their criteria, they didn't know how to evaluate and determine which was which and the program failed. But do you know what? The program failed two years ago and they still have the assistant principal. That's Parkinson's Law in operation.

But back to research and development. You ought to have research and development right at the high school. You have a principal and you may well have a very high-level person who is the executive officer or whatever you want to call him. The assistant principal or vice-principal in charge of operations runs the store and keeps it going well. Then you should have the other one, the assistant principal or vice-principal or whatever you want to call him, in charge of break-through, research and development, right there on that campus—the research person. There aren't enough of them I grant. We probably aren't training enough to do this all the way but I bet we are training enough who know how to play the Devil's Advocate and who, if we put them in the right spot in the organization with the right kind of power, could ask the uncomfortable question. Why do you do it that way? Because the best approach to responsible innovation is in having somebody play the Devil's Advocate.

In the traditional management sense this is called "decentralization." I believe that if we are going to have responsible change in typical schools, not in just the three or four would-be isolated schools throughout the United States, but responsible change going on every place at a pace fast enough to keep us from losing whatever it is we are working for, we are going to have to build fast. To do this we will have to take advantage of what I think is the most precious thing in the world, and that is the human being—his brain and his intelligence. We are getting better teachers all the time. Fountainheads must come up all over the place. We can no longer prescribe. Teachers are getting militant. Do you know why they

are getting militant? Because they are being treated paternally, like peons. They really are. It isn't really money they want so much as to be treated as professionals. It is touch and go as to whether or not they are going to be a professional or whether they are going to go another route. It is touch and go as to who is going to determine which way it is going to go. Our ability as administrators and staff people and psychologists is to help and to learn how to work with teachers to bring about positive change, right there at the grass-roots level of the operation. I feel so strongly about this that I guess I sound like a Baptist minister when I talk about it.

Let me get to one other idea. We are trying to work out something that, for want of a better word we are calling it a "development laboratory." It's a kind of a phenomenon in a school district. It's a place where, when somebody has an idea, somebody will be there to help them to put it into some kind of a plan or format to bring it to fruition. It's a place not to generate ideas and impose them on other people but to help teachers give birth to their own ideas. Too often a teacher who has an idea goes to his principal and says, "I have an idea I want to try out. I think it would be terrific. I think it would improve the efficiency of our operation." Do you know what they are asked? "Does it cost any money?" "Well, not much, but just a little bit." "Is it in the budget?" "No. No. I guess it isn't in the budget." "Well, then, I will tell you what you do. You just keep it hot and next spring we will see whether it can survive, when we start —." Now, you know how hot it will be. Teachers are busy people and that idea will get cold very fast.

What would happen, now, if right in your organization, in your budget, not as an adjunct, but right in the middle of it, you put the development lab where somebody can go when he gets a hot idea? Now, I don't want to make it easy. Make them stand the test, have a panel of interrogators or something, or a Devil's Advocate, poking holes and trying to show that the idea isn't any good. But, then, if it does have some merit, fund it with your "hot idea" fund.

Some people think that is ridiculous. However, it *will* work. We call it the "hot idea fund." Fifty thousand bucks. You know, people learn that the fund is there, but, then we have a problem. Do you know what? Teachers and principals don't know how—I shouldn't say this. I am not being negative about it when I say this. They are busy people. They are not supposed to know how to organize projects, but they can be helped. When you dangle $50,000 out

there, it's the dog-gonedest motivation. You know, it's a lot better motivation than saying, "Why don't you do something different in your classrooms?"

You dangle this and they want some of it. Then you set the conditions. You have the right to. The money is there and you have every right to set some conditions. The conditions are that they must have a plan and the plan must include objectives. The plan must include a plan to accomplish the objectives. And then it must include something else. "How are you going to know when you get through that it made a difference?" Now, those are simple words. Those aren't even statistical words are they? But the teachers understand those words and they understand them real well. They get excited about it.

I think my whole thesis—and I am practically through—is that you can organize to improve. It depends upon what you value and where you place that value. You can organize so that teachers and principals will work on responsible innovation if you dangle the carrots in the right places and if you establish a system to get it done.

Let me mention one more thing. It is a little rough to have a full time research and development plan with an assistant in every school at the elementary level. I believe in establishing the position of a "learning analyst" in each school. If our function primarily is learning and we provide for a principal and teachers, then this is somebody whose job is to help them analyze whether they are achieving their goals or not. Then you evaluate the principal and the teachers in terms of whether they achieve those goals. I think in this way you have some action going. You establish a climate for human dialogue, not cold statistics. You have somebody asking questions. A teacher recently reported about how excited the first grade teachers were because somebody asked and talked about what they were doing. They need this and they need somebody to put it in the kind of terms that will allow them to know when they get through whether it made any difference.

Now let me close by saying that, and I say this because I feel it strongly, with all the innovation and all of the things that we work on to improve education, there is just one thing we must never forget—the most important thing that goes on in any school occurs when a kid gets to feel good about himself. Whenever this happens (and I think you know what I mean) we have achieved. So, along with all of this you must continually work on making absolutely

sure that every kid that goes to school finds somebody who cares. This is more important, I think, than any organization of time or space or resources or anything else. And the innovation that we work on is really an attempt to try to work out a better way to accomplish this particular goal.

I wish we would quit talking about process all of the time. I will close with one little illustration. I am putting a lot of pressure now on trying to get some things going to increase efficiency, particularly at the high school level. A principal came to me the other day and said, "Well, I'm ready." And I asked, "You are ready for what?" And he said, "I am ready to go into flexible scheduling next fall." And he really expected me to say, "My God, man, that's great. You are really it." I said, "What in the heck do you want to do that for?" He was frustrated. He said, "What do you mean? You have been screaming around here about innovating and getting things done." And I said, "Well, now, what in the world would you want to go into flexible scheduling for?" He really didn't have a very good answer. I am sure that the real reason he wanted to go into flexible scheduling was because he thought I wanted him to. I said, "It's the last thing I want you to do, Buddy. Let's not go on flexible scheduling there unless it helps to meet our primary objective—learning." He walked out talking to himself.

It's going to take some time to get to this point where improving school learning is the target. But I repeat what I said originally—you need to create some problems in order to come along with some good solutions.

Conference List • 1967

Georgia S. Adams, *California State College*, Los Angeles, California
John D. Adams, Jr., *Burlingame High School*, Burlingame, California
Michael W. Adams, *Sacramento City Unified School District*, Sacramento, California
Frank Alt, *Turlock School District*, Turlock, California
Gerald J. Alves, *Sonoma State College*, Rohnert Park, California
Ray W. Alvord, *Stanford University*, Stanford, California
Robert D. Anderson, *Miramonte High School*, Orinda, California
Scarvia B. Anderson, *Educational Testing Service*, Princeton, New Jersey
Gene Atwood, *Tracy Joint Union High School*, Tracy, California
J. Douglas Ayers, *California Test Bureau*, Monterey, California
Alden W. Badal, *Oakland Unified School District*, Oakland, California
Michele Battermann, *Educational Testing Service*, Berkeley, California
Arthur G. Becker, *Stockton Unified School District*, Stockton, California
Robert Beebe, *Desert Sun School*, Idyllwild, California
Paul E. Bender, *Educational Testing Service*, Berkeley, California
Mary Betts, *Palo Alto Unified School District*, Palo Alto, California
John C. Bianchini, *Educational Testing Service*, Berkeley, California
Don F. Blood, *Western Washington State College*, Bellingham, Washington
Marilyn Blum, *Scripps College*, Claremont, California
Wayne Boone, *San Rafael High School*, San Rafael, California
Walter R. Borg, *Far West Laboratory*, Berkeley, California
Jack C. Bowen, *Napa County Schools*, Napa, California
Robert Bowman, *Watsonville High School*, Watsonville, California
Birdie N. Boyles, *California State Department of Education*, Sacramento, California
Donald Boysen, *Saratoga High School*, Saratoga, California
David J. Brinks, *University of Oregon Counseling Center*, Eugene, Oregon
Barbara M. Brown, *California Institute of Technology*, Pasadena, California
Frank R. Brown, *Nevada State Department of Education*, Carson City, Nevada
Guy H. Browning, *Stanford University*, Stanford, California
Katherine I. Brownlie, *Educational Testing Service*, Berkeley, California

Robert G. Cameron, *Western Regional Office, College Entrance Examination Board*, Palo Alto, California
Vincent N. Campbell, *American Institute for Research*, Palo Alto, California
Richard W. Carey, *Palo Alto Unified School District*, Palo Alto, California
Alberta Carlen, *Jefferson Union High School District*, Daly City, California
Lucille E. Carson, *George Washington High School*, San Francisco, California
Gerald P. Cavanaugh, *Denver Public Schools*, Denver, Colorado
Margaret Lee Chadwick, *Margaret Lee Chadwick Counseling Center*, Palos Verdes, California
A. Edward Chally, *Davis Joint Unified School District*, Davis, California
Thomas R. Citizen, *Westmont College*, Santa Barbara, California
R. W. Douglas Clack, *Educational Testing Service*, Berkeley, California
Edwin C. Clark, *Burbank Unified School District*, Burbank, California
James V. Clark, *Office of Relations with Schools, University of California*, Berkeley, California
Joan P. Cochran, *Educational Testing Service*, Berkeley, California
Pasquale P. Contestable, *John Swett High School*, Crockett, California
Merrill R. Convis, *Sir Francis Drake High School*, San Anselmo, California
Reginald A. Corder, *Educational Testing Service*, Berkeley, California
Mrs. Jean Corry, *Lee Senior Elementary School*, Woodland, California
Hugh W. Cox, *University of California*, Berkeley, California
K. Patricia Cross, *Educational Testing Service*, Berkeley, California
Walter F. Crowley, *Madera Unified School District*, Madera, California
Roger W. Cummings, *San Francisco State College*, San Francisco, California
Ferdinand Dagenais, *Educational Testing Service*, Berkeley, California
Thomas P. Davies, *San Jose Unified School District*, San Jose, California
John N. Davis, *San Jose Unified School District*, San Jose, California
Richard R. DeBlassie, *New Mexico State University*, Las Cruces, New Mexico
Frederick M. DeBruler, *Educational Testing Service, Northwest Regional Office*, Seattle, Washington
Ralph DeBruler, *Edmonds School District No. 15*, Lynnwood, Washington
Curt Demele, *Palo Alto Senior High School*, Palo Alto, California
Eileen Dombrow, *Ravenswood High School*, East Palo Alto, California
Glenn W. Durflinger, *School of Education, University of California*, Santa Barbara, California.
Nathan A. Edwards, *Department of Education, Fresno State College*, Fresno, California
Mrs. Richard W. Elliott, *Desert Sun School*, Idyllwild, California
Ralph F. Evans, *Department of Education, Fresno State College*, Fresno, California
Richard C. Falk, *Tacoma Community College*, Tacoma, Washington
Richard R. Fickel, *San Lorenzo Valley Unified School District*, Ben Lomond, California
John C. Flanagan, *American Institute for Research*, Palo Alto, California

Earl Flegel, *Multnomah County Intermediate Education District*, Portland, Oregon
Thomas Flom, *Science Research Associates, Inc.*, Oakland, California
Theodore T. Fong, *Alum Rock Union Elementary School District*, San Jose, California
Richard O. Fortna, *Educational Testing Service*, Berkeley, California
Mrs. Theodore E. Foster, San Francisco, California
Garford G. Gordon, *California Teachers Association*, Burlingame, California
Merrill V. Goudie, *Los Angeles County Schools*, Los Angeles, California
Mary Frances Gould, *Palo Alto Unified School District*, Palo Alto, California
Gordon R. Graves, *Fresno City Unified School District*, Fresno, California
John W. Grindy, *Calaveras County Office of Education*, San Andreas, California
Lois Guernsey, *Simpson Bible College*, San Francisco, California
Howard R. Haering, *San Bernardino County Schools*, San Bernardino, California
Edward A. Hakkarainen, *Oakland Unified School District*, Oakland, California
Betsey R. Haley, *Fremont Union High School District*, Sunnyvale, California
Clarence Hall, *Pacific Grove Unified School District*, Pacific Grove, California
Jan Handke, *School of Education, Stanford University*, Stanford, California
Lucille Harris, *Orange County Schools*, Santa Ana, California
J. Richard Harsh, *Southern California Office, Educational Testing Service*, Los Angeles, California
Einar Haugen, *Golden Gate Academy*, Oakland, California
Harrison F. Heath, *San Jose State College*, San Jose, California
John S. Helmick, *Educational Testing Service*, Princeton, New Jersey
Walter S. Hertzog, Jr., *East Los Angeles College*, Los Angeles, California
Thomas F. Humiston, *City College of San Francisco*, San Francisco, California
William Hutchinson, *Cubberley High School*, Palo Alto, California
William Jackman, *Holmes Senior High School*, Davis, California
Janet H. Jacobs, *La Sierra College*, Riverside, California
Ollie A. Jensen, *California State Department of Education*, Sacramento, California
Mrs. Iva Jex, *School Testing Service*, Berkeley, California
Robert A. Jones, *Test Bureau, University of Southern California*, Los Angeles, California
Elva Jovisich, *Western Regional Office, College Entrance Examination Board*, Palo Alto, California
Irene Jurgensen, *Lodi Union High School*, Lodi, California
Warren Kallenbach, *School of Education, San Jose State College*, San Jose, California

Philip G. Kearney, *Tulare County Department of Education*, Visalia, California
John G. Kelly, *Diablo Valley College*, Pleasant Hill, California
Evelyn B. Kipp, *Newark Unified School District*, Newark, California
Walter L. Klas, *Livermore Valley Unified Schools*, Livermore, California
Santelia A. Knight, *Educational Testing Service*, Berkeley, California
Nadine M. Lambert, *School of Education, University of California*, Berkeley, California
Robert E. Lambert, *Educational Testing Service*, Berkeley, California
Vernon F. Larsen, *Science Research Associates, Inc.*, Oakland, California
Curtis G. Larson, *San Mateo High School*, San Mateo, California
Alexander I. Law, *Evaluation and Research, State Department of Education*, Sacramento, California
Deloris J. Leander, *Educational Testing Service*, Berkeley, California
Carole A. Leland, *Western Regional Office, College Entrance Examination Board*, Palo Alto, California
Sister M. Leonard, *Presentation High School*, San Francisco, California
Charles Lewin, *San Luis Coastal Unified School District*, San Luis Obispo, California
Hazel M. Lewis, *University of the Pacific*, Stockton, California
John H. Lewis, *Cubberley High School*, Palo Alto, California
Brooks J. Lockhart, *Naval Postgraduate School*, Monterey, California
Carlos O. Lopez, *Pajaro Valley Unified School District*, Watsonville, California
Donald W. Lucas, *Oakland Technical High School*, Oakland, California
Clifford Earl Lunneborg, Jr., *Bureau of Testing, University of Washington*, Seattle, Washington
Donald M. McCauley, *Pleasant Hill High School*, Pleasant Hill, California
Cecil A. McDonald, *Vallejo Unified School District*, Vallejo, California
Benjamin W. McKendall, Jr., *Western Regional Office, College Entrance Examination Board*, Palo Alto, California
Arnold M. Martin, *Yerington High School*, Yerington, Nevada
Vera M. Martin, *Yerington Elementary School*, Yerington, Nevada
John L. Martin, *Palo Alto Unified School District*, Palo Alto, California
Eldon Matal, *Woodland High School*, Woodland, California
Chester O. Mathews, *University of California*, Santa Barbara, California
Chester L. Maul, *San Mateo County Schools*, Redwood City, California
Herbert G. May, *Santa Fe High School*, Santa Fe Springs, California
Alan Means, *Westmont College*, Santa Barbara, California
Arlene Metha, *Chapman College*, Orange, California
Robert G. Metters, *St. George's School*, Spokane, Washington
Richard C. Minard, *Educational Testing Service*, Berkeley, California
Anita M. Mitchell, *Culver City Unified School District*, Culver City, California
Charles C. Moffatt, *Armstrong College*, Berkeley, California
David B. Muirhead, *University of Colorado*, Boulder, Colorado
Mrs. Myra Nebel, *South San Francisco School District*, South San Francisco, California

James E. Nelson, *Western Regional Office, College Entrance Examination Board,* Palo Alto, California
Oberlin B. Nereson, *Foothill College,* Los Altos Hills, California
Leland B. Newcomer, *Newport-Mesa Unified School District,* Newport Beach, California
Gloria Nichols, *Redwood City School District,* Redwood City, California
James Norton, *San Rafael High School,* San Rafael, California
S. Gene Odle, *The College of Idaho,* Caldwell, Idaho
Stanley R. Ostrom, *Santa Clara County Office of Education,* San Jose, California
Mary H. Owen, *Educational Testing Service,* Berkeley, California
Paul D. Plowman, *Mentally Gifted, California State Department of Education,* Sacramento, California
Paul E. Plummer, *Golden Gate Academy,* Oakland, California
Gary S. Poor, *Pomona College,* Claremont, California
George V. Porter, Jr., *Examinations Division (CRE), United States Air Force Academy,* Colorado
Helen Prouty, *School of Education, San Diego State College,* San Diego, California
Mabel C. Purl, *Riverside Unified Schools,* Riverside, California
Gladys Quinlin, *Chadwick School,* Palos Verdes Estates, California
Richard M. Regedal, *Kern County Union High School District,* Bakersfield, California
Joy Richardson, *Oakland Public Schools,* Oakland, California
Christine M. Robison, *Katherine Delmar Burke School,* San Francisco, California
Thomas B. Rodgers, *Educational Testing Service,* Berkeley, California
David Rook, *Lincoln Elementary School,* Exeter, California
Susan Roosenberg, *Westmont College,* Santa Barbara, California
Dorothy C. Roy, *Educational Testing Service,* Berkeley, California
Glenn Royer, *Napa Valley Unified School District,* Napa, California
Rodger W. Russell, *Calaveras Unified School District,* San Andreas, California
David M. Sawyer, *Monterey Peninsula Unified School District,* Monterey, California
Gilbert Sax, *University of Washington,* Seattle, Washington
Wade W. Schroeder, *Campbell Union High School District,* Campbell, California
Bernard E. Seaman, *Northwest Nazarene College,* Nampa, Idaho
Alan J. Seder, *Educational Testing Service,* Berkeley, California
Jack B. Selbig, *University of Nevada,* Reno, Nevada
James Shannon, *Stockton Unified School District,* Stockton, California
Wesley Sharpe, *Livermore Unified School District,* Livermore, California
Catherine Sheehan, *George Washington High School,* San Francisco, California
Gerald R. Shelby, *Guidance and Pupil Personnel Services, Nevada State Department of Education,* Carson City, Nevada
Lee T. Sheldon, *Woodland Joint Unified School District,* Woodland, California

Thomas A. Shellhammer, *Bureau of Evaluation & Research, California State Department of Education,* Sacramento, California
Ralph P. Sherlock, *Merced Unified High School District,* Merced, California
Edmund O. Shinn, *College of San Mateo,* San Mateo, California
Leslie E. Shuck, *Newport-Mesa Unified School District,* Newport Beach, California
William G. Siemens, *Biola College,* La Mirada, California
Herbert S. Singer, *Berkeley Unified School District,* Berkeley, California
Rodney W. Skager, *School of Education, University of California,* Los Angeles, California
William M. Slaichert, *School of Education, University of Denver,* Colorado.
William R. Smith, *Sacramento County Office of Education,* Sacramento, California
Dorothy Soeberg, *Riverside County Schools,* Riverside, California
Sherry Sonneveldt, *Westmont College,* Santa Barbara, California
James E. Sudyk, *Gunn High School,* Palo Alto, California
Meribau Swank, *Educational Data Services, Palo Alto Unified School District,* Palo Alto, California
John Barrett Taylor, *California College of Arts and Crafts,* Oakland, California
Margaret L. Thomas, *Fresno City Unified Schools,* Fresno, California
Alice G. Thompson, *Department of Psychology, California State College,* Los Angeles, California
Anton Thompson, *Long Beach Unified School District,* Long Beach, California
Edward R. Tibby, *Educational Testing Service,* Berkeley, California
Frank Toggenburger, *Los Angeles City Schools, Research and Development,* Los Angeles, California
Edward Tonningsen, *Mt. Diablo Unified School District,* Concord, California
Pauline Toschi, *College of Marin,* Kentfield, California
R. Jack Totheroh, *Ventura County Schools,* Ventura, California
Joseph D. Tranchina, *Mission San Jose High School,* Fremont, California
Vernon E. Trimble, *Saratoga High School,* Saratoga, California
Walter Jack Vassar, *San Leandro High School,* San Leandro, California
Donald D. Vacks, *Newark Unified School District,* Newark, California
Bruce Wainwright, *Napa County Schools,* Napa, California
Annette C. Waite, *Educational Testing Service,* Berkeley, California
C. Eugene Walker, *Division of Education and Psychology, Westmont College,* Santa Barbara, California
Douglas R. Walker, *Stanford University,* Stanford, California
Harvey R. Wall, *Mt. Diablo Unified School District,* Concord, California
T. W. Walters, *Department of Education, North Pacific Union Conference, Seventh Day Adventists,* Portland, Oregon
Richard W. Watkins, *Educational Testing Service,* Berkeley, California
Leo Weisbender, *Los Angeles City Schools,* Los Angeles, California

John Wennergren, *Gonzales Union High School,* Gonzales, California
Charles L. Wheeler, Jr., *Fresno State College,* Fresno, California
June Whiteford, *Newark Unified School District,* Newark, California
Howard Wigell, *John Swett High School,* Crockett, California
Gertrude Williams, *Davis Senior High School,* Davis, California
Larry Williams, *Exeter High School,* Exeter, California
Margaret C. Williams, *Yreka Union High School,* Yreka, California
Thomas Williams, *Richmond Schools,* Richmond, California
Barbara Wilson, *Woodland Joint Unified High School District,* Woodland, California
Peter F. Wolmut, *Multnomah County Intermediate Education District,* Portland, Oregon
Robert G. Wolstoncroft, *Sierra High School, Whittier Union High School District,* Whittier, California
Deborah Wubben, *University of California,* Santa Cruz, California
Carl R. Wulfsberg, *Long Beach Unified School District,* Long Beach, California
Harold Zindell, *Ravenswood High School,* East Palo Alto, California